

This Page Is Inserted by IFW Operations  
and is not a part of the Official Record

## **BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning documents *will not* correct images,  
please do not report the images to the  
Image Problem Mailbox.**

(19) RÉPUBLIQUE FRANÇAISE  
INSTITUT NATIONAL  
DE LA PROPRIÉTÉ INDUSTRIELLE  
PARIS

(11) N° de publication :  
(à n'utiliser que pour les  
commandes de reproduction)

2 715 937

(21) N° d'enregistrement national :

94 01532

(51) Int Cl<sup>6</sup> : C 12 N 7/00, C 12 Q 1/68, A 61 K 31/70, 38/16, C 07 H  
21/00

(12)

## DEMANDE DE BREVET D'INVENTION

A1

(22) Date de dépôt : 04.02.94.

(71) Demandeur(s) : Société Anonyme dite: BIO  
MERIEUX — FR.

(30) Priorité :

(43) Date de la mise à disposition du public de la  
demande : 11.08.95 Bulletin 95/32.

(72) Inventeur(s) : Perron Hervé, Mallet François,  
Mandrand Bernard, Bédin Frédéric et Beseme  
Frédéric.

(56) Liste des documents cités dans le rapport de  
recherche préliminaire : Ce dernier n'a pas été  
établi à la date de publication de la demande.

(73) Titulaire(s) :

(60) Références à d'autres documents nationaux  
apparentés :

(74) Mandataire : Cabinet Germain & Maureau.

(54) Virus MSRV1 et MSRV2 associés ensemble à la sclérose en plaques et leurs constituants nucléiques.

(57) L'invention concerne l'association de deux virus asso-  
ciés à la sclérose en plaques, à savoir, un premier virus hu-  
main, possédant une activité transcriptase inverse, et asso-  
cié à une famille d'éléments rétroviraux endogènes, ou un  
variant du premier virus comprenant au moins un antigène  
reconnu par au moins un anticorps dirigé contre au moins  
un antigène correspondant dudit premier virus, et un se-  
cond virus, possédant une activité transcriptase inverse, ou  
un variant dudit second virus, ces deux virus étant issus  
ensemble d'une même souche virale choisie parmi les sou-  
ches dénommées respectivement POL-2, déposée le  
22.07.1992 auprès de l'ECACC sous le numéro d'accès  
V92072202 et MS7PG déposée le 08.01.93 auprès de  
l'ECACC sous le numéro d'accès V93010816, et parmi les  
souches variantes consistant en des virus comprenant au  
moins un antigène qui est reconnu par au moins un anti-  
corps dirigé contre au moins un antigène correspondant de  
l'un ou l'autre des virus des souches virales POL-2 et  
MS7PG précitées.

FR 2 715 937 - A1



**Virus MSRV1 et MSRV2 associés ensemble à la Sclérose en Plaques et leurs constituants nucléiques**

La Sclérose en Plaques (SEP) est une maladie démyélinisante du système nerveux central (SNC) dont la 5 cause reste encore inconnue.

De nombreux travaux ont étayé l'hypothèse d'une étiologie virale de la maladie, mais aucun des virus connus testés ne s'est avéré être l'agent causal recherché: une revue des virus recherchés depuis des 10 années dans la SEP a été faite par E. Norrby (Prog. Med. Virol., 1978; 24, 1-39) et R.T. Johnson(dans "Handbook of clinical neurology, 47 Demyelinating diseases". Vinken P. et Bruyn G.W., eds. Amsterdam, Elsevier science Publishing, 1985, 319-336). Parallèlement, la possibilité 15 d'un facteur exogène et/ou infectieux est suggérée par l'existence d'épidémies localisées ou "clusters" de SEP comme ce qui a été observé dans les Iles Feroes entre 1943 et 1960 (Cook 1980), en Sardaigne (Rosati, 1988), en Norvège (Riisse, 1991), ainsi que par les études sur les 20 migrants (Elian, 1990). Parmi tous les facteurs exogènes suggérés, les virus ont été étudiés le plus souvent et une étiologie virale est classiquement évoquée.

L'observation, dans la SEP, de phénomènes assimilables à une réaction d'auto-immunité a conduit à 25 une hypothèse étiologique auto-immune "essentielle" (voir: Lisak R.P., Zweiman B. New Engl. J. Med. 1977; 297, 850-853, et, Lassmann H. et Wisniewski H.M. Arch. Neurol. 1979; 36, 490-497). Cependant, cette auto-immunité dirigée contre certains composants du SNC s'est révélée peu 30 spécifique de la SEP et fréquente dans les inflammations du SNC, associées ou non à une infection, ainsi que cela a été montré par Hirayama M. et coll. (Neurology 1986; 36, 276-8), Kenneth G. Warren et coll. (Annals of Neurology 1986; 20, 20-25), Suzumura A. et coll. (Journal of 35 Neuroimmunology 1986; 11, 137-47), et, Tourtelotte W. et coll. (Journal of Neurochemistry 1986; 46, 1086-93). De

plus, comme l'a fait remarquer E.J. Field (The Lancet 1989; I, 1272.) aucune des thérapeutiques immuno-suppressives n'a permis d'obtenir des résultats décisifs contre la SEP. Il semble maintenant probable que les 5 manifestations "auto-immunes" sont induites par un mécanisme d'origine virale : co-sensibilisation à des déterminants vitaux associés à des molécules d'origine cellulaire, phénomènes de mimétisme moléculaire -comme cela a été décrit par Fujinami R.S et Oldstone M.B.A 10 (dans "Molecular mimicry: Cross-reactivity between microbes and host proteins as a cause of autoimmunity". Oldstone M.B.A., ed. Current Topics in Microbiology and Immunology, vol. 145, Berlin, Springer-Verlag, 1989)- ou, selon P. Rudge (Journal of Neurology, Neurosurgery and 15 Psychiatry, 1991 54, 853-855), par expression de superantigènes rétroviraux .

Des travaux ont étayé une hypothèse selon laquelle un Rétrovirus serait à l'origine de la maladie : la découverte récente par A. Gessain et coll. (J. Infect. 20 Disease 1988; 1226-1234), de syndromes neurologiques associés au virus HTLV-I, connu à l'origine comme agent de leucémies T de l'adulte, a conduit de nombreux auteurs, tels que H. Koprowski et coll. (Nature 1985; 318, 154), M.Ohta et coll. (J. Immunol. 1986; 137, 3440), E. P. Reddy 25 et coll. (Science 1989; 243, 529), S.J. Greenberg et coll. (Proc. Natl. Acad. Sci. USA 1989; 86, 2878), J.H. Richardson et coll. (Science 1989; 246, 821), S.L. Hauser et coll. (Nature 1986; 322, 176) et A. Karpas et coll. (Nature 1986; 322, 177), à rechercher une implication de 30 ce rétrovirus humain dans la SEP, cependant sans succès ou avec des résultats évoquant des réactions croisées.

Par ailleurs, il existe un modèle animal très proche de la SEP, induit par un rétrovirus : le virus MAEDI-VISNA du mouton. Il est connu que l'infection 35 naturelle par ce virus peut provoquer deux types de maladies chez cet animal: le Maedi, une pneumonie

interstitielle lymphocytaire qui peut survenir lors de la primo-infection par ce rétrovirus et une pathologie neurologique démyélinisante tardive suivant généralement une phase de latence prolongée, le Visna. La physiopathologie du Visna naturel concorde avec la plupart des données cliniques et biologiques de la SEP, comme le rapportent Johnson R.T. (Rev. Infect. Dis. 1985; 7, 66-67), Narayan O. et Cork L.C. (Rev. Infect. Dis. 1985; 7, 89-98), et Nathanson N. et coll. (Rev. Infect. Dis. 1985; 7, 75-82). L'infection expérimentale des moutons par inoculation intra-ventriculaire de souches neurovirulentes du virus VISNA a permis d'établir la responsabilité de ce virus dans la génèse de cette affection démyélinisante du mouton. Comme l'expliquent Nathanson N. et coll. (Rev. Infect. Dis. 1985; 7, 75-82), Hoffman P.M. et Panitch H.S. (dans, "Handbook of clinical neurology, 12 : Viral diseases" R.R. Mc Kendall, ed. Elsevier science Publishing, Amsterdam, 1989, p 453-466), et A. Haase (Nature 1986; 322, 130-136), elle diffère de l'infection naturelle par ses conséquences neuropathologiques exacerbées, mais reste proche de la SEP. Il est de plus notable que, dans l'ensemble des travaux effectués sur ce sujet par les auteurs précités, notamment, le virus Visna est régulièrement retrouvé dans les cellules de plexus choroïdes du cerveau qui constituent un site de latence et de réplication occasionnelle du provirus Visna; la localisation de ces cellules à l'interface sang/liquide céphalo-rachidien (LCR) explique certainement ce phénomène.

De plus, le modèle du Visna apporte une donnée supplémentaire pour l'approche rétrovirologique de ce type de maladie. En effet, il s'avère que, dans le cadre de l'infection naturelle de ces moutons, des coïncidences épidémiologiques et biologiques ont été observées entre d'une part les pathologies du Maedi-Visna et d'autre part, l'adénomatose ovine pulmonaire ou Jaagsiekte, pathologie tumorale provoquée par un autre rétrovirus, le JSRV (De

Martini J.C. et coll. J.N.C.I. 1987; 79, 167-177 / Rosadio R.H. et coll. Vet. Pathol. 1988; 25, 58-66 / Dawson M. et coll. Br. Vet. J. 1990; 146, 531-537). Le JSRV est un rétrovirus endogène, voire la souche exogène associée à 5 une famille d'éléments endogènes, dont les propriétés infectieuses et oncogènes ont été mises en évidence (York D. F. et coll. J. Virol. 1992; 66, 4930-4939). Or, il est notable que, dans des isolats naturels de moutons affectés de maladies attribuées soit au rétrovirus Maedi-Visna, 10 soit au rétrovirus Jaagsiekte, on retrouve généralement les deux rétrovirus (Dawson M. et coll. Br. Vet. J. 1990; 146, 531-537 / York D.F. et coll. J. Virol. 1991; 65, 5061-5067). Cette coexpression et/ou coinfection semble indiquer un rôle synergique de ces deux rétrovirus dans le 15 cadre du Maedi (De Martini J.C. et coll. J.N.C.I. 1987; 79, 167-177). Ainsi, dans certaines techniques de purification d'isolats naturels sur gradient de saccharose on a pu mettre en évidence l'existence de deux pics de sédimentation séparés correspondant respectivement au JSRV 20 et au rétrovirus Maedi-Visna (York D.F. et coll. J. Virol. 1991; 65, 5061-5067). Ces données permettent d'entrevoir la complexité d'une approche analytique visant à caractériser un agent rétroviral associé à une maladie donnée, à partir d'isolats naturels dans lesquels peuvent 25 coexister plusieurs agents pathogènes.

Récemment, un rétrovirus, différent des rétrovirus humains connus a été isolé chez des patients atteints de SEP par H. Perron et coll. (Res. Virol. 1989; 140, 551-561/ dans : "Current concepts in multiple sclerosis" 30 Wiethölter et coll., eds. Amsterdam, Elsevier, 1991, p. 111-116 / The Lancet 1991; 337, 862-863 ). Les auteurs ont aussi pu montrer que ce rétrovirus pouvait être transmis in vitro, que des patients atteints de SEP produisaient des anticorps susceptibles de reconnaître des protéines 35 associées à l'infection des cellules leptoméningées par ce rétrovirus et que l'expression de ce dernier pouvait être

fortement stimulée par les gènes immédiats-précoces de certains herpesvirus (Perron et coll. Herpes simplex virus ICP0 and ICP4 immediate-early proteins strongly enhance expression of a retrovirus harboured by a leptomeningeal 5 cell-line from multiple sclerosis. 1993. J. Gen. Virol. 74; 65-72).

Tous ces résultats plaident en faveur du rôle dans la SEP d' au moins un rétrovirus inconnu ou d'un virus ayant une activité transcriptase inverse détectable selon 10 la méthode publiée par H. Perron et coll. (Res. Virol. 1989; 140, 551-561) et qualifiée d'activité "RT de type LM7".

Récemment, les travaux de la demanderesse ont permis d'obtenir deux lignées continues de cellules 15 infectées par des isolats naturels provenant de deux patients différents atteints de SEP, par un procédé de culture tel que décrit dans la demande de brevet n° WO 93/20188. Ces deux lignées dérivées de cellules de plexus-choroïdes humains, dénommées LM7PC et PLI-2 ont été 20 déposées à l'E.C.A.C.C. respectivement le 22 juillet 1992 et le 8 janvier 1993, sous les numéros 92072201 et 93010817, conformément aux dispositions du traité de Budapest. Par ailleurs, les isolats viraux possédant une activité RT de type LM7 ont également été déposés à 25 1'E.C.A.C.C. sous la dénomination globale de "souches". La "souche" ou isolat hébergé par la lignée PLI-2, dénommée POL-2, a été déposée le 22 juillet 1992 sous le n° V92072202. La "souche" ou isolat hébergé par la lignée LM7PC, dénommée MS7PG, et a été déposée le 8 janvier 1993 30 sous le n° V93010816.

A partir des cultures et des isolats précités, caractérisés par des critères biologiques et morphologiques, on s'est ensuite attaché à caractériser le matériel nucléique associé aux particules virales 35 produites dans ces cultures.

Ainsi, les objets de l'invention sont les suivants:

- une association de deux virus associés à la sclérose en plaques, à savoir, un premier virus humain, 5 possédant une activité transcriptase inverse, et associé à une famille d'éléments rétroviraux endogènes, ou un variant du premier virus comprenant au moins un antigène reconnu par au moins un anticorps dirigé contre au moins un antigène correspondant dudit premier virus, et un 10 second virus, possédant une activité transcriptase inverse, ou un variant dudit second virus, ces deux virus étant issus ensemble d'une même souche virale choisie parmi les souches dénommées respectivement POL-2, déposée le 22.07.1992 auprès de l'ECACC sous le numéro d'accès 15 V92072202 et MS7PG déposée le 08.01.93 auprès de l'ECACC sous le numéro d'accès V93010816, et parmi les souches variantes consistant en des virus comprenant au moins un antigène qui est reconnu par au moins un anticorps dirigé contre au moins un antigène correspondant de l'un ou 20 l'autre des virus des souches virales POL-2 et MS7PG précitées;

- une association de deux virus associés à la sclérose en plaques, à savoir, un premier virus humain possédant une activité transcriptase inverse, et associé à 25 une famille d'éléments rétroviraux endogènes, ou un variant du premier virus comprenant au moins un antigène reconnu par au moins un anticorps dirigé contre au moins un antigène correspondant dudit premier virus, et un second virus possédant une activité transcriptase inverse, 30 ou un variant du second virus, ces deux virus étant produit par une même lignée cellulaire choisie parmi les souches dénommées respectivement PLI-2 déposée le 22.07.1992 auprès de l'ECACC sous le numéro d'accès 92072201 et LM7PC déposée le 08.01.93 auprès de l'ECACC 35 sous le numéro d'accès 93010817, et par toutes cultures cellulaires infectées susceptibles de produire un virus

comportant au moins un antigène qui est reconnu par au moins un anticorps dirigé contre au moins un antigène correspondant de l'un ou l'autre des virus produits par les lignées PLI-2 et LM7PC précitées;

5 - une association de deux virus, à savoir un premier virus dont le génome comprend une séquence nucléotidique présentant au moins 50 % et préférentiellement au moins 70 % d'homologie avec la séquence nucléotidique SEQ ID N05 ou sa séquence  
10 complémentaire, SEQ ID N05 comportant mises bout à bout, SEQ ID N01, SEQ ID N03 et SEQ ID N04; SEQ ID N01, SEQ ID N03 et SEQ ID N04 étant décrites respectivement à la Fig 1, à la Fig 9 et à la Fig 20, ou un variant dudit premier virus dont le génome code pour une séquence  
15 peptidique présentant au moins 50 % et préférentiellement au moins 70 % d'homologie avec la séquence peptidique codée par la séquence nucléotidique SEQ ID N01, décrite à la Fig 1, ou sa séquence complémentaire, et un second virus, dont le génome comprend une séquence nucléotidique  
20 présentant au moins 50 % et préférentiellement au moins 70 % d'homologie avec une séquence nucléotidique choisie parmi SEQ ID N02, décrite à la Fig 19, ou sa séquence complémentaire, ou un variant dudit second virus dont le génome code pour une séquence peptidique présentant au  
25 moins 50 %, et préférentiellement au moins 70 % d'homologie avec la séquence peptidique codée par la séquence nucléotidique SEQ ID N02 ou sa séquence complémentaire;

30 - un procédé de détection d'un premier virus, et/ou d'un second virus, associés à la sclérose en plaques, caractérisé en ce qu'on met en oeuvre au moins deux fragments, un premier fragment dont la séquence nucléotidique présente au moins 50 % et de préférence au moins 70 % d'homologie avec SEQ ID N05 ou sa séquence  
35 complémentaire, et/ou un second fragment dont la séquence nucléotidique présente au moins 50 % et de préférence au

moins 70 % d'homologie avec SEQ ID N02 ou sa séquence complémentaire, chacun desdits fragments étant une sonde ou une amorce;

5 - une composition diagnostique, prophylactique ou thérapeutique, caractérisée en ce qu'elle comprend une association de l'invention et un premier fragment nucléotidique, et/ou un second fragment nucléotidique, tels que définis dans l'invention;

10 - un procédé pour détecter et/ou identifier une association de virus associés à la sclérose en plaques, dans un échantillon biologique, caractérisé en ce qu'on expose un ARN et/ou un ADN spécifique à chaque dit virus, et/ou leur ARN et/ou ADN complémentaire à une association d'un premier fragment nucléotidique et d'un second 15 fragment nucléotidique, tels que définis dans l'invention;

20 - un procédé de détection d'un premier virus, et/ou d'un second virus, associés à la sclérose en plaques, caractérisé en qu'on met en oeuvre un premier peptide codé de manière partielle ou totale par le premier fragment nucléotidique défini dans l'invention, et/ou un second peptide codé de manière partielle ou totale par le deuxième fragment nucléotidique défini dans l'invention;

25 - une composition diagnostique, prophylactique ou thérapeutique, caractérisée en ce qu'elle comprend un premier peptide et/ou un second peptide, définis dans l'invention;

30 - une composition diagnostique, prophylactique ou thérapeutique, caractérisée en ce qu'elle est obtenue par préparation d'un premier ligand spécifique du premier peptide, et/ou un second ligand spécifique du second peptide, et par association du premier ligand et/ou du second ligand;

35 - un fragment nucléotidique, caractérisé en ce que sa séquence nucléotidique présente au moins 50% , et de préférence au moins 70% d'homologie, avec une séquence choisie parmi SEQ ID N01, SEQ ID N03, SEQ ID N04, et la

séquence SEQ ID N05 comprenant bout à bout SEQ ID N01, SEQ ID N03, SEQ ID N04.

Avant de détailler l'invention, différents termes 5 utilisés dans la description et les revendications sont à présent définis :

- selon l'invention, un fragment nucléotidique ou un oligonucléotide est un enchaînement de monomères, caractérisé par la séquence informationnelle des acides 10 nucléiques naturels, susceptibles de s'hybrider à un fragment nucléotidique dans des conditions prédéterminées, l'enchaînement pouvant contenir des monomères de structures différentes et être obtenu à partir d'une molécule d'acide nucléique naturelle et/ou par 15 recombinaison génétique et/ou par synthèse chimique,

- ainsi un monomère peut être un nucléotide naturel d'acide nucléique dont les éléments constitutifs sont un sucre, un groupement phosphate et une base azotée; dans l'ARN le sucre est le ribose, dans l'ADN le sucre est 20 le désoxy-2-ribose; selon qu'il s'agisse de l'ADN ou l'ARN, la base azotée est choisie parmi l'adénine, la guanine, l'uracile, la cytosine, la thymine; ou un nucléotide modifié dans l'un au moins des trois éléments constitutifs; à titre d'exemple, la modification peut 25 intervenir au niveau des bases, générant des bases modifiées telles que l'inosine, la méthyl-5-désoxycytidine, la désoxyuridine, la diméthylamino-5-désoxyuridine, la diamino-2,6-purine, la bromo-5-désoxyuridine et toute autre base modifiée favorisant 30 l'hybridation, au niveau du sucre à savoir le remplacement d'au moins un désoxyribose par un polyamide (P.E. Nielsen et al, Science, 254, 1497-1500 (1991)), au niveau du groupement phosphate par exemple son remplacement par des esters notamment choisis parmi les esters de diphosphate, 35 d'alkyl et arylphosphonate et de phosphorothioate,

- par "séquence informationnelle", on entend toute suite ordonnée de monomères, dont la nature chimique et l'ordre dans un sens de référence constituent une information de même qualité que celle des acides nucléiques naturels,
- par hybridation, on entend le processus au cours duquel, dans des conditions appropriées, deux fragments nucléotidiques, ayant des séquences suffisamment complémentaires, s'apparient pour former un double brin,
- une sonde est un fragment nucléotidique comprenant de 10 à 100 monomères, avantageusement de 10 à 50 monomères, possédant une spécificité d'hybridation dans des conditions déterminées pour former un complexe d'hybridation avec un fragment nucléotidique ayant une séquence nucléotidique comprise dans l'ARN du virus de l'invention, l'ADN obtenu par transcription inverse dudit ARN et l'ADN dont ledit ARN est le produit de transcription; une sonde peut être utilisée à des fins de diagnostic telles que les sondes de capture et/ou de détection ou à des fins de thérapie,
- la sonde de capture peut être immobilisée sur un support solide par tout moyen approprié, c'est-à-dire directement ou indirectement, par exemple par covalence ou adsorption passive,
- la sonde de détection est marquée au moyen d'un marqueur choisi parmi les isotopes radioactifs, des enzymes notamment choisis parmi la peroxydase et la phosphatase alcaline et ceux susceptibles d'hydrolyser un substrat chromogène, fluorogène ou luminescent, des composés chimiques chromophores, des composés chromogènes, fluorogènes ou luminescents, des analogues de bases nucléotidiques, et la biotine,
- les sondes utilisées à des fins de diagnostic de l'invention peuvent être mises en oeuvre dans toutes les techniques d'hybridation connues et notamment les techniques dites "DOT-BLOT". (MANIATIS et al, Molecular

Cloning, Cold Spring Harbor, 1982), "SOUTHERN BLOT" (SOUTHERN. E.M., J. Mol. Biol., 98, 503 (1975), "NORTHEN BLOT" qui est une technique identique à la technique "SOUTHERN BLOT" mais qui utilise de l'ARN comme cible, la 5 technique SANDWICH (DUNN A.R., HASSEL J.A., Cell, 12, 23 (1977)); avantageusement, on utilise la technique SANDWICH dans la présente invention comprenant une sonde de capture 10 spécifique et/ou une sonde de détection spécifique, étant entendu que la sonde de capture et la sonde de détection doivent présenter une séquence nucléotidique au moins partiellement différente,

- une autre application de l'invention est une sonde de thérapie antisens, ladite sonde étant susceptible de s'hybrider *in vivo* sur l'ARN et/ou sur l'ADN pour 15 bloquer les phénomènes de traduction et/ou de transcription,

- une amorce est une sonde comprenant de 10 à 30 monomères, possédant une spécificité d'hybridation dans des conditions déterminées pour l'initiation d'une 20 polymérisation enzymatique par exemple dans une technique d'amplification telle que la PCR (Polymerase Chain Reaction), dans un procédé d'elongation, tel que le séquençage, dans une méthode de transcription inverse ou analogue,

25 - l'homologie caractérise le degré de similitude de deux fragments nucléotidiques comparés.

Etant donné qu'un virus possédant une activité enzymatique transcriptase inverse peut être génétiquement caractérisé aussi bien sous forme d'ARN que d'ADN, il sera 30 fait mention aussi de l'ADN que de l'ARN viral pour caractériser les séquences relatives à un virus possédant une telle activité transcriptase inverse.

L'invention sera mieux comprise à la lecture de la description détaillée qui va suivre faite en référence aux 35 figures annexées dans lesquelles:

- la figure 1, représente la séquence en acides nucléiques du clone PSJ17, (SEQ ID N01),
- la figure 2, donne un alignement de la séquence nucléique PSJ17 avec celle d'un rétrovirus connu ERV9,
- 5 - la figure 3, correspond à la séquence en acides aminés de la séquence PSJ17,
- la figure 4, donne la comparaison des séquences protéiques de SPJ17 et ERV9 théorique (selon La Mantia et col. (1991) ),
- 10 - la figure 5 représente les séquences des différents clones MSRV-1B obtenus après PCR dans la région "pol" définie par Shih et col.,
- la figure 6, représente l'arbre phylogénique des séquences MSRV-1B obtenues par PCR dans la région "pol" 15 définie par Shih et col.,
- la figure 7, donne la définition d'une trame de lecture fonctionnelle pour chaque type de famille de type MSRV-1B/"PCR pol",
- la figure 8, représente, selon Fig 8A, l'alignement des séquences protéiques déduites des consensus A, 20 B, C et D de type MSRV-1B avec des séquences d'origine rétrovirale, selon Fig 8B l'arbre phylogénique et selon Fig 8C, l'homologie par rapport à ERV9,
- la figure 9 représente les séquences consensus 25 de type MSRV-1B obtenues à partir d'ARN viral, issu des lignées LM7PC et PLI-2, amplifié par PCR dans la région "pol", ainsi qu'un consensus général comprenant les amorces d'amplification, et dénommé SEQ ID N03,
- la figure 10 est une représentation de 30 l'activité transcriptase inverse (RT) en dpm (désintégration par minute) dans les fractions de saccharose prélevées sur un gradient de purification des virions produits par les lymphocytes B en culture d'un patient atteint de SEP.
- 35 - la figure 11 donne dans les mêmes conditions qu'à la figure 10 le dosage de l'activité transcriptase

inverse dans la culture d'une lignée B obtenue à partir d'un témoin exempt de sclérose en plaques ,

5 - la figure 12 représente, selon Fig 12A, les séquences consensus et majoritaires de type MSRV-1B obtenues à partir des lymphocytes B d'un patient atteint de SEP, avec

10 \* MAJ correspondant à la séquence majoritaire dans laquelle les bases conservées dans tous les cas sont représentées par ATGC, les bases majoritaires sont représentées par atgc et les bases délétées par rapport au consensus sont représentées par - ,

15 \* VAR correspondant à la séquence majoritaire (variation) dans laquelle les bases conservées par rapport au consensus sont représentées par ., les bases minoritaires sont représentées par atgc et les bases délétées par rapport au consensus sont représentées par - et

20 \* MIN correspondant aux séquences minoritaires (exception) dans lesquelles les bases conservées par rapport à la séquence majoritaire sont représentées par . et les bases délétées par rapport au consensus par - ,

ainsi que, selon Fig 12B, la traduction de la séquence majoritaire, avec

25 an: acides nucléiques:

- en gras, caractères de petites tailles: les sites de restriction des amorces
- en gras, caractères majuscules, extrémité 3' des amorces
- en souligné, la séquence nucléique majoritaire codante

prot: séquence protéique:

- en souligné: acides aminés codés par les amorces

- la figure 13 représente la séquence de type MSRV-2A obtenue à partir des cultures LM7 selon le protocole priginal de Shih et col.,

5 - la figure 14, représente l'alignement d'une séquence protéique MSRV-2A obtenue à partir de la culture LM7 avec des séquences protéiques rétrovirales connues; les acides aminés conservés entre le produit d'amplification (Trad pol SHIH) et certains rétrovirus sont soulignés dans l'alignement. Cette séquence est très 10 éloignée de tous les rétrovirus connus à ce jour, en particulier de HTLV, notamment par la présence d'une délétion dans la région 5',

15 - la figure 15, donne un exemple d'une séquence consensus pour les séquences de type MSRV-2B,

15 - la figure 16 représente la définition d'une séquence consensus de type MSRV-2B selon Fig 16A, une définition d'une trame de lecture fonctionnelle selon Fig 16B, une répétition de l'amorce 3' selon Fig 16C et une séquence comprise entre VLPQG et YVDD selon Fig 16D,

20 - la figure 17, représente l'alignement d'une séquence protéique MSRV-2B (avec 1 ou 2 amorces 3') avec des séquences protéiques rétrovirales connues; les acides aminés conservés (1ère ligne) entre le consensus MSRV-2B et certains rétrovirus sont soulignés dans l'alignement.

25 Cette séquence est très éloignée de tous les rétrovirus connus à ce jour, en particulier de HTLV, notamment par la présence d'une délétion dans la région 5' et d'un site actif YVDD et non YMDD,

30 - la figure 18, représente les séquences consensus et majoritaires de MSRV-2B obtenues à partir des lymphocytes B d'un patient atteint de SEP, selon Fig 18A,

35 \* MAJ correspondant à la séquence majoritaire dans laquelle les bases conservées dans tous les cas sont représentées par ATGC, les bases majoritaires sont représentées par atgc, et

\* VAR correspondant à la séquence majoritaire (variation) dans laquelle les bases minoritaires sont représentées par atgc  
 ainsi que, la traduction de la séquence majoritaire, selon  
 5 Fig 18B, dans laquelle la légende est la suivante :

an: acides nucléiques:

- \* en gras, caractères de petites tailles: les sites de restriction des amorces
- \* en gras, caractères majuscules, extrémité 3' des amorces
- \* en souligné, la séquence nucléique majoritaire codante

prot: séquence protéique:

\* en souligné: acides aminés codés par les amorces  
 15 - la figure 19, donne une analyse comparée des séquences nucléiques (Fig 19A) et protéiques (Fig 19B) de type MSRV-2 obtenues à partir des cultures LM7, LM7PC et PLI-2 et des lymphocytes B d'un patient atteint de SEP, selon le protocole original ou modifié de Shih et Col., la 20 séquence consensus étant dénommée SEQ ID N02, et avec LigT4 correspondant à une séquence obtenue à partir des lignées, selon le protocole modifié, LBpatDuT4 correspondant à une séquence obtenue à partir de patient, selon le protocole modifié et polSHIH correspondant à une 25 séquence obtenue à partir des lignées, selon le protocole original,

- la figure 20, représente l'alignement du consensus MSRV-1B, avec le produit de PCR "nichée" M003-P004, tel que défini dans l'exemple 8, et le clone PSJ17  
 30 (MSRV-1A) avec la légende suivante :

\* en souligné:

amorce M003-BCD telle que décrite dans l'exemple 8

amorce P004-BCD telle que décrite dans l'exemple 8

\* en gras:

35 séquence intermédiaire entre PSJ17 et MSVR-1B

et montre que les séquences MSRV-1A et B appartiennent à la même famille virale et vraisemblablement à la même souche virale et constituent deux parties différentes d'un même génome viral, elle 5 représente en outre la séquence intermédiaire PSJ17 et MSRV1-B, dénommée SEQ ID N04.

**EXEMPLE 1: OBTENTION D'UN CLONE PSJ17, DEFINISSANT  
UNE FAMILLE MSRV-1A, PAR REACTION DE TRANSCRIPTASE INVERSE  
10 ENDOGENE SUR UNE PREPARATION DE VIRION ISSUE DE LA LIGNEE  
PLI-2.**

Cette approche vise à obtenir des séquences d'ADN rétrotranscrites à partir de l'ARN supposé rétroviral dans l'isolat en utilisant l'activité transcriptase inverse 15 présente dans ce même isolat. Cette activité transcriptase inverse ne peut théoriquement fonctionner qu'en présence d'un ARN rétroviral, lié à un ARNt amorce ou hybridé à des brins courts d'ADN déjà rétro-transcrits dans les particules rétrovirales (Lori F. et coll. J. Virol. 1992; 20 66,5067-5074). Ainsi l'obtention de séquences rétrovirales spécifiques dans un matériel contaminé par des acides nucléiques cellulaires, était optimisée grâce à l'amplification enzymatique spécifique des portions d'ARN viraux par une activité transcriptase inverse virale. Les 25 auteurs ont, pour cela, déterminé les conditions physico-chimiques particulières dans lesquelles cette activité enzymatique de transcription inverse sur des ARN contenus dans des virions pouvait être effective in vitro; ceux-ci correspondent à la description technique des protocoles 30 présentés ci-dessous (réaction de RT endogène, purification, clonage et séquençage).

L'approche moléculaire a consisté à utiliser une préparation de virion concentré, mais non purifié, obtenu à partir des surnageants de culture de la lignée PLI-2 35 préparé selon la méthode suivante: les surnageants de cultures sont collectés deux fois par semaine, pré-

centrifugés à 10 000 tpm pendant 30 minutes pour éliminer les débris cellulaires, et ensuite congelés à -80°C ou utilisés tels que pour les étapes suivantes. Les surnageants frais ou décongelés sont centrifugés sur un 5 coussin de PBS-glycérol 30% à 100 000g (ou 30 000 tpm dans un rotor LKB-HITACHI de type 45 T) pendant 2h à 4°C. Après élimination du surnageant, le culot sédimenté est repris dans un petit volume de PBS et constitue la fraction de virions concentrés, mais non purifiés. Cet échantillon 10 viral concentré mais non purifié a été utilisé pour effectuer une réaction dite de transcription inverse endogène, telle que décrite ci-après: un volume de 200µl de virions purifiés selon le protocole décrit ci-dessus et contenant une activité transcriptase inverse d'environ 1-5 15 millions de DPM est décongelé à 37°C jusqu'à apparition d'une phase liquide, puis placé sur de la glace. Un tampon 5 fois concentré a été préparé avec les composants suivants: Tris-HCl pH 8.2, 500 mM; NaCl 75 mM; MgCl<sub>2</sub> 25 mM; DTT 75 mM et NP 40 0.10 %. 100 µl de tampon 5X + 25 µl 20 d'une solution de dATP 100mM + 25 µl d'une solution de dTTP 100mM + 25 µl d'une solution de dGTP 100mM + 25 µl d'une solution de dCTP 100mM + 100 µl d'eau distillée stérile + 200 µl de la suspension de virions (activité T.I. de 5 million de DPM) dans le PBS ont été mélangés et 25 incubés à 42°C pendant 3 heures. Après cette incubation le mélange réactionnel est directement mélangé avec un mélange tamponné phénol/chloroforme/alcool isoamylque (Sigma ref. P 3803), la phase aqueuse est collectée et un volume d'eau distillée stérile est ajouté à la phase 30 organique pour ré-extraire le matériel nucléique résiduel. Les phases aqueuses collectées sont regroupées et les acides nucléiques contenus sont précipités par addition d'acéate de sodium 3M pH 5,2 au 1/10 de volume + 2 volumes d'éthanol + 1µl de glycogène (Boehringer-Mannheim 35 ref. 901 393) et mise à -20°C de l'échantillon pendant 4h ou la nuit à +4°C. Le précipité obtenu après

centrifugation est ensuite lavé avec de l'éthanol à 70% et resuspendu dans 60 ml d'eau distillée. Les produits de cette réaction ont ensuite été purifiés, clonés et séquencés selon le protocole décrit ci-après: des ADN

5 bouts francs avec des adénines non-appariées aux extrémités ont été générés: une réaction de "remplissage" a d'abord été effectuée: 25  $\mu$ l de la solution d'ADN précédemment purifiée ont été mélangés avec: 2 $\mu$ l d'une solution 2,5 mM contenant, en quantité équimolaire, dATP +

10 dGTP + dTTP + dCTP / 1  $\mu$ l d'ADN polymerase T4 (Boehringer-Mannheim ref. 1004 786) / 5  $\mu$ l de 10X "incubation buffer for restriction enzyme" (Boehringer-Mannheim ref. 1417 975) / 1  $\mu$ l d'une solution à 1% de sérum-albumine bovine / 16  $\mu$ l d'eau distillée stérile. Ce mélange a été incubé 20

15 minutes à 11°C. 50 $\mu$ l de tampon TE et 1 $\mu$ l de glycogène (Boehringer-Mannheim ref. 901 393) y ont été ajoutés avant extraction des acides nucléiques avec du phénol/chloroforme/alcool isoamylique (Sigma ref. P 3803), et précipitation à l'acétate de sodium comme décrit

20 précédemment. L'ADN précipité après centrifugation est resuspendu dans 10  $\mu$ l de tampon 10 mM Tris pH 7.5. Puis, 5 $\mu$ l de cette suspension ont été mélangés avec 20 $\mu$ l de tampon Taq 5X, 20  $\mu$ l de 5mM dATP, 1 $\mu$ l (5U) de Taq ADN polymerase (AmpliTaq<sup>TM</sup>) et 54  $\mu$ l d'eau distillée stérile.

25 Ce mélange est incubé 2 h à 75°C avec un film d'huile à la surface de la solution. L'ADN en suspension dans la solution aqueuse prélevée en dessous du film d'huile après incubation est précipité comme décrit précédemment et resuspendu dans 2  $\mu$ l d'eau distillée stérile. L'ADN obtenu

30 a été inséré dans un plasmide à l'aide du kit TA Cloning<sup>TM</sup>. Les 2 $\mu$ l de solution d'ADN ont été mélangés avec 5 $\mu$ l d'eau distillée stérile, 1 $\mu$ l d'un tampon de ligation 10 fois concentré "10X LIGATION BUFFER", 2 $\mu$ l de "pCRT<sup>TM</sup> VECTOR" (25 ng/ml) et 1 $\mu$ l de "TA DNA LIGASE". Ce mélange a

35 été incubé une nuit à 12°C. Les étapes suivantes ont été réalisées conformément aux instructions du kit TA Cloning<sup>®</sup>

(British Biotechnology). A la fin de la procédure, les colonies blanches de bactéries recombinantes (white) ont été repiquées pour être cultivées et permettre l'extraction des plasmides incorporés selon la procédure 5 dite de "miniprep" (Sambrook J., Fritsch E.F., et Maniatis T., Molecular cloning, a laboratory manual. Cold spring harbor laboratory press, 1989). La préparation de plasmide de chaque colonie recombinante a été coupée par une enzyme de restriction appropriée et analysée sur gel d'agarose.

10 Les plasmides possédant un insert détecté sous lumière UV après marquage du gel au bromure d'éthidium, ont été sélectionné pour le séquençage de l'insert, après hybridation avec une amorce complémentaire du promoteur Sp6 présent sur le plasmide de clonage du TA cloning kit.

15 La réaction préalable au séquençage a ensuite été effectuée selon la méthode préconisée pour l'utilisation du kit de séquençage "Prism ready reaction kit dye deoxyterminator cycle sequencing kit" (Applied Biosystems, ref. 401384) et le séquençage automatique a été réalisé 20 avec l'appareil "automatic sequencer, modèle 373 A Applied Biosystems selon les instructions du fabricant.

L'analyse discriminante sur les banques de données informatiques des séquences clonées à partir des fragments d'ADN présents dans le mélange réactionnel a permis de 25 mettre en évidence trois catégories de séquences: la première correspondait à des séquences totalement inconnues dans les banques de données et sans aucune parentée avec des gènes décrits. La seconde, correspondait à des séquences totalement homologues à des portions de 30 gènes cellulaires connus. La troisième correspondait à des séquences présentant des homologies partielles plus ou moins importantes avec des séquences connues de rétrovirus ou d'éléments génétiques rétrotransposables comme les rétroposons ou rétrotransposons.

35 L'origine des séquences de la première catégorie est difficile à définir alors que pour la seconde, elle

s'explique par la présence attendue de fragments d'ADN cellulaire contaminant la fraction d'acides nucléiques extrait de l'échantillon viral concentré mais non-purifié, après incubation dans les conditions requises pour la 5 réaction de transcription inverse endogène sus-mentionnée.

L'origine des séquences de la troisième catégorie est manifestement rétroviral ou de type rétroviral. Il est connu que des séquences rétrovirales endogènes apparentées à un rétrovirus réplicatif, ou co-exprimées dans une même 10 cellule infectée, peuvent être encapsidées dans les virions générés par une souche rétrovirale donnée (Linial M.L. and Miller A.D. dans "Current topics in microbiology and immunobiology. Retroviruses, strategies of replication" vol. 157, p. 125-152; Swanstrom R. et Vogt P.K. éditeurs, 15 Springer-Verlag, Heidelberg 1990) et la présence de certaines séquences de type rétroviral peut s'expliquer par ce phénomène.

Cependant, parmi les séquences de cette catégorie, une séquence a été retrouvée le plus fréquemment et s'est 20 avérée représenter la majorité des clones obtenus dans ces conditions à partir de l'isolat viral "POL-2" issu de la culture de la lignée PLI-2 entre le quarantième et le soixantième passage en culture après établissement de la lignée continue selon le principe décrit dans le brevet 25 n°WO-93/20188 au nom de la Demandante. Le clone correspondant PSJ17 a été entièrement séquencé et la séquence obtenue, présentée dans la figure 1, a été analysée à l'aide du logiciel "Geneworks®" sur les banques de données actualisées "Genebank®". L'analyse des banques 30 de données n'a pas permis de trouver de séquence identique déjà décrite. Seule une homologie partielle a pu être retrouvée avec certains éléments rétroviraux connus. L'homologie relative la plus intéressante concerne un rétrovirus endogène dénommé ERV-9, ou HSERV-9, selon les 35 références (La Mantia et coll. Nucleic Acids Research 1991;19, 1513-1520). Ainsi la séquence du clone PSJ17 a pu

etre localisee dans la region pol codant pour la transcriptase inverse d'un rétrovirus, ou encore d'un virus possédant une enzyme avec une activité transcriptase inverse, par alignement avec la séquence décrite pour 5 HSERV-9, tel que cela est présenté présenté dans la figure 2. La traduction en acides aminés du clone PSJ17 est présentée dans la figure 3, et la comparaison avec cette séquence protéique et la séquence protéique théorique de ERV9 telle que décrite par La Mantia et col est présentée 10 dans la figure 4.

**EXEMPLE 2: OBTENTION DE CLONES APPARENTES A PSJ17,  
DEFINISSANT UNE FAMILLE MSRV-1A, PAR AMPLIFICATION  
ENZYMATIQUE (RT-PCR) SUR UNE PREPARATION DE VIRION ISSUE  
15 DE LA LIGNEE PLI-2.**

Des amorces nucléiques, PSJ17A et PSJ17B, ont été définies dans la séquence du clone PSJ17 et ont été utilisées pour rechercher la présence de l'ARN et/ou de 20 l'ADN correspondant à un virus contenant cette séquence, dans différents échantillons biologiques. Celles-ci sont définies ci-après:

Séquence amorce sens PSJ17 A : AAATTCCTCACTACCTG

Séquence amorce antisens PSJ17 B : GGGTTTAAGAGTTGCACA

25

Les réactions de PCR ont été faites dans un volume total de 100  $\mu$ l, contenant 200ng d'ADN, 33  $\mu$ mole de chaque primer, 0.25mM de chaque dNTP, 10 $\mu$ l de tampon 10X (10 fois concentré) et 2,5  $\mu$ l d'enzyme Taq. Les cycles 30 d'amplification sont réalisés comme suit: dénaturation 95°C/5 minutes, hybridation des amorces 50°C à 55°C/2 minutes, extension 72°C/2,5 minutes, puis pendant 33 cycles, 95°C/2minutes, 50°C à 55°C/2 minutes, 72°C/2,5 minutes, et enfin 95°C/2minutes, 50°C à 55°C/2 minutes et 35 72°C/7 minutes.

Les réactions de RT-PCR, selon un un procédé d'amplification d'ARN tel que décrit dans la demande de brevet n° EP 0 569 272 A1, ont été faites dans un volume total de 100  $\mu$ l, contenant 200ng d'ARN, 1  $\mu$ l de RNA Guard, 5 33  $\mu$ mole de chaque primer, 0.25mM de chaque dNTP, 10 $\mu$ l de tampon 10X, 2,5 U d'enzyme Taq. et 0,4  $\mu$ l d'enzyme RT (10u) sont aussi ajoutés aux échantillons. Les cycles d'amplification sont réalisés comme suit: dénaturation de l'ARN 65°C/10 minutes, synthèse de l'ADNC 50°C/8 minutes 10 puis les cycles sont identiques à ceux de la PCR décrite ci-dessus. Des réactions de contrôle ont été effectuées de manière à contrôler l'absence de contaminants (réaction sur de l'eau). Les produits ont été analysés sur gel d'acrylamide 10%.

15 Une des propriétés de la polymérase Taq consistant à ajouter une adénine à l'extrémité 3' de chacun des deux brins d'ADN, les produits issus de la PCR et de la RT-PCR ont été purifiés selon le procédé décrit dans l'exemple 1 puis directement clonés à l'aide du kit TA Cloning® 20 (British Biotechnology), puis séquencés selon le protocole décrit dans l'exemple 1.

Le matériel biologique utilisé consiste en virion purifié à partir de l'isolat "POL-2" (ECACC n° V92072202) selon la procédure décrite ci-après: les surnageants de 25 cultures sont collectés deux fois par semaine, pré-centrifugés à 10 000 tpm pendant 30 minutes pour éliminer les débris cellulaires, et ensuite congelés à -80°C ou utilisés tels que pour les étapes suivantes. Les surnageants frais ou décongelés sont centrifugés sur un 30 coussin de PBS-glycérol 30% à 100 000g (ou 30 000 tpm dans un rotor LKB-HITACHI de type 45 T) pendant 2h à 4°C. Après élimination du surnageant, le culot sédimenté est repris dans un petit volume de PBS et constitue la fraction de virions concentrés, mais non purifiés. Le virus concentré 35 est ensuite déposé sur un gradient de sucre dans un tampon PBS stérile (15 à 50% poids/poids) et

ultracentrifugé à 35 000 tpm (100 000g) pendant 12 h à +4°C dans un rotor à godets. 10 fractions sont recueillies et 20µl sont prélevés dans chaque fraction après homogénéisation pour y doser l'activité transcriptase 5 inverse selon la technique décrite par H. Perron et coll. (Res. Virol. 1989; 140, 551-561). Les fractions contenant le pic d'activité RT de type LM7 sont ensuite diluées dans du tampon PBS stérile et ultracentrifugées une heure à 35 000 tpm (1000 000 g) pour sédimenter les particules 10 virales. Le culot de virions purifiés ainsi obtenu est alors repris dans un petit volume d'un tampon adéquat pour l'utilisation ultérieure qui en sera faite (Ex. Tampon Guanidium Thiocyanate pour l'extraction des ARN; PBS stérile pour le stockage à -80°C).

15 Notamment, un procédé d'amplification d'ARN tel que décrit dans la demande de brevet n° EP 0 569 272 A1, utilisé sur de l'ARN purifié, utilisant des amorces telles que décrites ci-dessus, a permis d'amplifier du matériel nucléique sur une fraction de virion purifié sur 20 gradient de saccharose selon la méthode décrite ci-dessus à partir de l'isolat "POL-2" (ECACC n° V92072202). Ce matériel nucléique une fois cloné et séquencé selon les techniques décrites dans l'exemple 1, s'est révélé être, dans ces virions purifiés, le même matériel génétique que 25 celui du clone PSJ17 obtenu à partir de l'échantillon de virions concentrés mais non-purifiés contenus dans un isolat "POL-2". Une réaction PCR avec les mêmes amorces sur le même matériel, mais sans étape de synthèse d'un ADNc à partir de l'ARN présent dans le matériel nucléique 30 de l'échantillon analysé, n'a permis aucune amplification de matériel nucléique. Ces résultats indiquent clairement que la séquence du clone PSJ17 est présente sous forme d'ARN et pas d'ADN, dans les fractions contenant le pic d'activité transcriptase inverse de type LM7, après 35 purification sur gradient de saccharose. Ces résultats sont compatibles avec la présence de particules virales

contenant un ARN génomique incluant la séquence du clone PSJ17 et une enzyme transcriptase inverse, liées à la présence d'un rétrovirus non caractérisé par ses séquences spécifiques jusqu'alors, dans les fractions purifiées de 5 virions produits par la lignée PLI-2.

Par ailleurs, la réaction RT-PCR sus-mentionnée (demande de brevet EP 0 569 272 A1) a été appliquée sur des échantillons biologiques, issus de patients témoins exempts de SEP, obtenus dans des conditions identiques 10 aux fractions de gradient de saccharose sus-mentionnées, à la différence près qu'aucune activité transcriptase inverse spécifique de type "LM7" n'y a été détectée. Les fractions de même densité en saccharose que les pics d'activité transcriptase inverse répertoriés (environ 15 1.17g/ml) ont ainsi été utilisées. Dans ces échantillons provenant respectivement d'une culture de cellules leptoméningées de patiente atteinte de maladie de Strumpell-Lorrain, d'une culture de plexus choroïdes d'une patiente décédée d'accident vasculaire cérébral, d'une 20 lignée de lymphocytes B (lignée lymphoblastoïde spontanée) d'un patient atteint de maladie génétique non précisée, tous étant exempts de SEP, cette réaction RT-PCR avec les amorces contenues dans la région "PSJ17" n'a pas permis d'amplifier de matériel nucléique ou, quand une 25 amplification a été visualisée dans un cas, le matériel amplifié s'est révélé n'avoir aucun rapport avec la séquence "PSJ17" (homologie partielle des amorces avec des gènes cellulaires). Ainsi, dans ces échantillons issus de témoins non-SEP et provenant de fractions de saccharose 30 homologues à celles qui contiennent des particules virales associées à une activité transcriptase inverse dans les échantillons "POL-2/PLI-2" et "MS7PG/LM7PC" provenant de SEP, cette séquence rétrovirale n'a pas été retrouvée associée à de l'ARN. Ces résultats sont compatibles avec 35 l'absence de production de particules rétrovirales

contenant la séquence du clone PSJ17 dans ce type de culture de cellules provenant de patients exempts de SEP.

Par contre ces séquences ont pu être détectées par cette même technique PCR, telle que décrite ci-avant, dans 5 des échantillons biologiques tels que des cellules humaines diverses, des sérumus obtenus après formation du caillot sanguin, des liquides céphalo-rachidiens obtenus par ponction lombaire et pouvant contenir des débris cellulaires, où de l'ADN humain était présent. Cela, chez 10 des patients atteints de SEP, mais aussi chez des patients témoins et des témoins sains. Une étude en southern-blot avec le clone PSJ17 utilisé comme sonde, a confirmé que la séquence du clone PSJ17 était apparentée à une famille de rétrovirus endogènes présents dans le génome humain. Etant 15 donnée la relative homologie avec le rétrovirus endogène ERV-9, il se peut que le rétrovirus contenant la séquence PSJ17 soit un membre d'une famille de rétrovirus endogènes apparenté à ERV-9, ou d'une nouvelle famille de rétrovirus endogènes proches d'ERV-9, ou encore, une souche exogène à 20 laquelle la famille endogène ERV-9 serait apparentée, dans la mesure où aucune homologie avec un rétrovirus exogène connu n'a été identifiée pour ERV-9, contrairement à toutes les séquences rétrovirales endogènes humaines caractérisées à ce jour, tel que cela apparaît dans le 25 tableau 1.

Ces résultats sont donc compatibles avec une expression d'ARN issu d'une famille de rétrovirus endogènes que la demanderesse a dénommé MSRV-1A, ayant une homologie relative avec ERV-9, mais contenant un consensus 30 de séquences nucléiques avec le clone de référence PSJ17. Ce virus MSRV-1A peut être exprimé de manière différente entre les patients atteints de SEP et les personnes exemptes de SEP, par tout type de cellules susceptible d'exprimer ce rétrovirus. Ces résultats sont aussi 35 compatibles avec la présence d'un rétrovirus exogène MSRV-1A, ayant une homologie relative avec ERV-9, mais

contenant un consensus de séquences nucléiques avec le clone de référence PSJ17, hébergé et exprimé par certaines cellules infectées chez les patients atteints de SEP et pas chez les témoins sains ou exempts de SEP.

5 La séquence obtenue dans le clone PSJ17 et sus-décrise, sera le mieux utilisée pour la détection stricte d'ARN spécifiquement exprimé chez les patients atteints de SEP et, conséquemment, de peptides ou protéines contenant la séquence d'acides aminés codée par cette séquence ARN

10 ou un consensus apparenté à cette dernière et provenant d'un variant MSRV-1A. De même l'utilisation de ces séquences peptidiques est conséquemment envisageable pour obtenir des antigènes qui permettront de détecter la réponse immunitaire à ces motifs antigéniques exprimés *in vivo*

15 chez les patients atteints de SEP, dans la mesure où les ARN correspondants sont exprimés *in vivo* et permettent la synthèse protéique correspondante, de manière originale dans la SEP ou tout autre pathologie individualisée et trouvée associée à ce même phénomène d'expression

20 rétrovirale.

**EXEMPLE 3: OBTENTION DE CLONES, DEFINISSANT UNE FAMILLE MSRV-1B, PAR AMPLIFICATION PCR "NICHEE" DES REGIONS POL CONSERVÉES DES RETROVIRUS SUR DES PRÉPARATIONS 25 DE VIRIONS ISSUES DES LIGNEES LM7 ET PLI-2.**

Une technique PCR dérivée de la technique publiée par Shih et coll. a été utilisée. Cette technique permet, par traitement de tous les composants du milieu réactionnel par la DNase, d'éliminer toute trace d'ADN contaminant. Elle permet parallèlement, par l'utilisation d'amorces différentes mais chevauchantes dans deux séries successives de cycles d'amplifications PCR, d'augmenter les chances d'amplifier un ADNc résultant d'une quantité 30 d'ARN faible au départ et encore réduite dans l'échantillon par l'action parasite de la DNase sur l'ARN;

35

ce, d'autant plus la DNase est utilisée dans des conditions d'activité en excès qui permettent d'éliminer toute trace d'ADN contaminant avant inactivation de cette enzyme restant dans l'échantillon par chauffage à 85°C 5 pendant 10 minutes. Cette variante de la technique PCR décrite par Shih et coll, a été utilisée sur des fractions de virions, purifiés comme précédemment pour les surnageants congelés des cultures LM7, issus cette fois de lignée lymphoblastoïde B spontanée provenant d'un nouveau 10 cas de SEP, de la culture PLI-2 (ECACC n° 92072201) et de la culture LM7PC (ECACC n° 93010817), ces deux dernières cultures ayant été obtenues selon les méthodes ayant fait l'objet du brevet n° WO 93/20188.

Après clonage avec le TA cloning kit® des produits 15 amplifiés par cette technique et analyse de la séquence à l'aide du séquenceur automatique selon ce qui est décrit dans les exemples 1 et 2, les séquences ont été analysées à l'aide du logiciel Geneworks®, sur la dernière version (n° 79, novembre 93) à ce jour de la banque de données 20 Genebank®.

Cette technique a d'abord été appliquée à deux fractions de virions purifiés comme décrit dans l'exemple 2 sur saccharose à partir de l'isolat "POL-2" produit par la lignée PLI-2 d'une part, et à partir de l'isolat MS7PG 25 produit par la lignée LM7PC d'autre part. Les séquences clonées et séquencées à partir de ces deux échantillons correspondent à quatre catégories. Une première catégorie de séquence, correspond à des séquences amplifiées à partir de matériel nucléique rétroviral contaminant la 30 transcriptase inverse du MoMuLV utilisée pour l'étape de synthèse de l'ADN complémentaire, ainsi que cela a déjà été signalé précédemment. La deuxième catégorie, retrouvée dans la majorité des clones (55% des clones issus des isolats POL-2 des culture PLI-2, et, 67% des clones issus 35 des isolats MS7PG des cultures LM7PC) correspond à une famille de séquences "pol" proches mais différentes du

rétrovirus endogène humain dénommé ERV-9 ou HSERV-9. La troisième catégorie correspond à des séquences très fortement homologues à la séquence attribuée à un virus nommé par la demanderesse MSRV-2B et décrit dans les 5 exemples 6 et 7. La quatrième catégorie correspond à diverses séquences trouvées généralement en un exemplaire, ainsi qu'une de type rétroviral endogène, cependant jamais retrouvées par d'autres approches ou dans d'autres échantillons provenant de culture exprimant une activité 10 transcriptase inverse de type LM7.

La deuxième catégorie de séquences représentant la majorité des clones est constituée de séquences dont la variabilité permet de définir quatre sous-familles de séquences. Ces sous-familles sont suffisamment proches 15 entre elles pour qu'on puisse les considérer comme des quasi-espèces provenant d'un même rétrovirus, tel que cela est bien connu pour le rétrovirus HIV-1 par exemple (Meyerhans et al., Cell 1989; 58, 901-910), ou comme le résultat de l'expression de plusieurs provirus endogènes 20 co-régulés dans les cellules productrices, puisqu'appartenant à la même famille de rétrovirus endogène et donc sensibles aux mêmes signaux de régulation générés éventuellement par le provirus répliquatif (Linial M.L. and Miller A.D. dans "Current topics in microbiology 25 and immunobiology. Retroviruses, strategies of replication" vol. 157, p. 125-152; Swanstrom R. et Vogt P.K. éditeurs, Springer-Verlag, Heidelberg 1990). Cette nouvelle famille de rétrovirus endogènes ou, alternativement, cette nouvelle espèce rétrovirale dont 30 nous avons obtenu en culture la génération de quasi-espèces, et qui contient un consensus des séquences décrites à la suite est dénommée MSRV-1B par la demanderesse.

Dans la figure 5 sont présentées les séquences des différents clones MSRV-1B séquencés lors de cette 35 expérience. Ces séquences présentent une homologie en acides nucléiques allant de 70% à 88% avec la séquence

HSERV9 référencée X57147 et M37638 dans la base de données Genebank. L'arbre phylogénétique de ces séquences est présenté dans la figure 6. Dans cette figure, les sous-familles A, B, C D représentent les séquences qui ont été 5 retrouvées de manière prépondérante dans des expériences similaires répétées ultérieurement, dans les échantillons d'ARN pur de virions purifiés avec les isolats MS7PG et POL-2. A partir de ces familles de séquences, quatre 10 séquences nucléiques "consensus" représentatives de différentes quasi-espèces d'un rétrovirus MSRV-1B éventuellement exogène ou de différentes sous-familles d'un rétrovirus endogène MSRV-1B, ont été définies. Ces 15 consensus représentatifs sont présentés dans la figure 7, avec la traduction en acides aminés. Une trame de lecture fonctionnelle existe pour chaque sous-famille de ces 20 séquences MSRV-1B et l'on peut voir que la trame de lecture ouverte fonctionnelle correspond à chaque fois à la séquence acides aminés venant en deuxième ligne sous la séquence acide nucléiques. Un alignement des consensus 25 A,B,C,D et d'un consensus général traduits en protéine avec les séquences de rétrovirus connus est présenté dans la figure 8. Le consensus général de la séquence MSRV-1B obtenue par cette technique PCR dans la région "pol" est présenté dans la figure 9.

25

**EXEMPLE 4: OBTENTION DE CLONES, DEFINISSANT UNE FAMILLE MSRV-1B, PAR AMPLIFICATION PCR "NICHEE" DES REGIONS POL CONSERVÉES DES RETROVIRUS SUR DES PRÉPARATIONS DE LYMPHOCYTES B D'UN NOUVEAU CAS DE SEP.**

30

La même technique PCR modifiée d'après la technique de Shih et coll. a été utilisée pour amplifier et séquencer le matériel nucléique ARN présent dans une fraction de virions purifiés au pic d'activité 35 transcriptase inverse "de type LM7" sur gradient de saccharose, selon les protocoles décrits dans l' exemple

2, à partir d'une lignée lymphoblastoïde spontanée, obtenue par auto-immortalisation en culture de lymphocytes B d'un patient SEP séropositif pour le virus d'Epstein-Barr (EBV) après mise en culture des cellules lymphoïdes 5 sanguine dans un milieu de culture approprié contenant une concentration appropriée de cyclosporine A. Une représentation de l'activité transcriptase inverse dans les fractions de saccharose prélevées sur un gradient de purification des virions produits par cette lignée (Duc.) 10 est présentée dans la figure 10. De même, les surnageants de culture d'une lignée B obtenue dans les mêmes conditions à partir d'un témoin exempt de sclérose en plaques ont été traités dans les mêmes conditions et le dosage de l'activité transcriptase inverse dans les 15 fractions du gradient de saccharose s'est avéré négatif partout (bruit de fond) et est présenté dans la figure 11. La fraction 3 du gradient correspondant à la lignée B de SEP et la même fraction sans activité transcriptase inverse du gradient témoin non-SEP, ont été analysées par 20 la même technique RT-PCR que précédemment, dérivée de Shih et coll., suivie des mêmes étapes de clonage et de séquençage tels que décrites dans les exemples 1 et 2.

L'analyse des clones recombinants prélevés au hasard a fourni les mêmes quatres catégories de séquences 25 que précédemment. La première catégorie (type 1) correspond à l'amplification de matériel rétroviral contaminant la transcriptase inverse de MoMuLV. La seconde catégorie (type 2) correspond aux séquences de type MSRV-1B telles que définies dans l'exemple 3. La troisième 30 catégorie correspond à des séquences très fortement homologues à la séquence attribuée à un virus nommé par la demanderesse MSRV-2B et décrit dans les exemples 6 et 7. Les résultats sont présentés dans le tableau 2. Il est tout à fait notable que les séquences de type MSRV-1B 35 soient retrouvées dans le seul matériel associé à un pic d'activité transcriptase inverse "de type LM7" provenant

de la lignée lymphoblastoïde B de SEP. Ces séquences n'ont pas été retrouvées avec le matériel de la lignée lymphoblastoïde B témoin (non-SEP) dans 26 clones recombinants pris au hasard. Seules les séquences 5 contaminantes de types Mo-MuLV et des séquences sans analogie rétrovirale particulière ont été retrouvées chez ce témoin. La différence de résultats est à l'évidence hautement significative (chi-2, p<0,001). L'analyse des séquences de type MSRV-1B obtenues à partir de la lignée B 10 de ce patient SEP est présentée dans la figure 12.

**EXEMPLE 5: OBTENTION DE CLONES, DEFINISSANT UNE FAMILLE MSRV-2A, PAR AMPLIFICATION DES REGIONS POL CONSERVEES DES RETROVIRUS SUR UNE PREPARATION DE VIRION 15 ISSUE DE LA LIGNEE LM7.**

L'approche moléculaire a consisté à utiliser une technique PCR permettant d'amplifier une région relativement conservée du gène pol des rétrovirus exogènes 20 et endogènes, mais aussi des virus codant pour une enzyme à activité transcriptase inverse tels que notamment le virus de l'hépatite B, qui a été publiée par Shih et coll. (Shih A., Misra R., and Rush M.G. Detection of multiple, novel reverse transcriptase coding sequences in human 25 nucleic acids : relation to primate retroviruses. J. Virol. 1989; 63, 64-75). Cette technique a été utilisée sur l'ARN extrait d'une préparation de virions purifiés obtenus selon le protocole décrit dans l'exemple 1, à partir des surnageants de la culture LM7 d'origine (Perron 30 et coll. Res. Virol. 1989; 140, 551-561) gardés congelés à -80°C depuis lors. Préalablement à la réaction PCR, l'ARN de l'échantillon a été transcrit en ADN complémentaire (ADNc) à l'aide du Kit "cDNA synthesis system plus" (Amersham), selon les instructions du fabricant et en se 35 basant sur une valeur approximée à un log près de la quantité d'ARN présente dans notre échantillon.

Après amplification PCR selon la technique de shih et coll., le matériel nucléique obtenu a été déposé sur un gel avec 2% d'agarose et la bande visualisée sous lumière ultraviolette après marquage au bromure d'éthidium dans 5 une région de poids moléculaire d'environ 100 paires de bases a été découpée et les acides nucléiques contenus, extraits selon le protocole usuel (Sambrook J., Fritsch E.F., et Maniatis T., Molecular cloning, a laboratory manual. Cold spring harbor laboratory press, 1989). puis 10 clonés en utilisant le kit TA cloning KIT® tel que décrit dans les exemples 1 et 2 (British Biotechnology) et les procédures conseillées dans les protocoles joints au kit. Après étalement des bactéries transformées sur le milieu de pousse approprié au protocoles du kit, quelques 15 dizaines de clones recombinants ont été repiqués et amplifiés afin d'obtenir une quantité suffisante de plasmide selon la technique dite de "miniprep" décrite dans Sambrook J., Fritsch E.F., et Maniatis T. (Molecular cloning, a laboratory manual. Cold spring harbor 20 laboratory press, 1989). L'insert des plasmides ainsi purifiés a été séquencé par séquençage automatique selon le protocole mentionné dans l'exemple 1.

Les séquences obtenues ont ensuite été analysées à l'aide des logiciels Mac Vector® et Geneworks® sur banque 25 de données informatiques Genebank®, pour les séquences nucléiques, et Swiss Prot®, pour les séquences en acides aminés déduites des trames de lecture mises en évidences dans les séquences nucléiques. L'analyse des séquences obtenues à partir d'échantillon viral provenant des 30 surnageants LM7 décongelés et purifié au pic d'activité transcriptase inverse sur gradient de saccharose, a mis en évidence trois catégories de séquences. Une première catégorie, correspondant à des associations artéfactuelles des amorces nucléiques utilisées pour l'amplification PCR, 35 une deuxième catégorie correspondant à des séquences sans trame de lecture ouverte consistante et une troisième

catégorie correspondant à des séquences rétrovirales dans la région "pol" attendue de par les amorces utilisées, et présentant au moins une trame de lecture ouverte conséquente.

5        Dans cette troisième catégorie, on a pu distinguer des séquences strictement homologues à des souches du murine leukaemia virus qui ont été retrouvées par la même approche méthodologique, dans des tubes témoins ne mettant en présence que de l'eau et l'enzyme transcriptase inverse  
10      10 du Moloney-murine leukaemia virus utilisée pour la synthèse de l'ADNc. Il est apparu évident que ces séquences provenaient de contaminants nucléiques associés à cette enzyme de rétrovirus murin utilisée pour cette étape réactionnelle. On a pu aussi y distinguer des  
15      15 séquences représentées par un nombre de clones allant de un à trois, et strictement homologues à des rétrovirus endogènes humains connus. Ces mêmes séquences ont été détectées dans des échantillons témoins ne provenant pas de SEP, traités dans les mêmes conditions. Enfin, on a  
20      20 aussi pu y distinguer une séquence représentée par une population majoritaire de clones (environ 42% des clones), relativement à la représentativité individuelle des autres séquences (toujours inférieure à 5%, voire 10% pour un petit nombre), et présentant des homologies partielles  
25      25 avec des rétrovirus connus dans la région "pol" attendue. L'interrogation de la banque de données Genebank® actualisée à ce jour (version 79, novembre 93) n'a pas permis de mettre en évidence de séquence identique ou présentant des homologies significatives.

30        Cette séquence est présentée dans la figure 13. Elle présente un cadre de lecture ouvert en phase avec les deux amorces PCR retrouvées aux extrémités mais elle est plus courte que l'ensemble des séquences rétrovirales connues dans la région attendue entre ces amorces. Une  
35      35 "délétion" de 45 paires de bases (15 acides aminés) y est observée à la suite de la séquence de l'amorce "amont",

comme cela est présenté dans la figure 14, alors que les séquences précédant l'amorce "aval" sont présentes. Cependant la trame de lecture est ouverte et ininterrompue sur toute la séquence incluant les amorces et la séquence 5 en acides aminés déduite présente une homologie significative avec la région correspondante des rétrovirus connus, comme cela est présenté dans la figure 14 . Dans la séquence interne aux amorces PCR, les acides aminés E, R, Q, P et D, normalement assez bien conservés dans cette 10 région pol des rétrovirus et des virus avec activité transcriptase inverse connus (shih et coll. J. Virol. 1989; 63, 64-75) sont retrouvés conservés aux bonne positions dans la trame de lecture de notre séquence originale.

15 Enfin, étant donné que cette séquence est suffisamment divergente des séquences rétrovirales déjà décrites dans les banques de données on peut avancer qu'il s'agit d'une séquence appartenant à un nouveau virus que la demanderesse a dénommé MSRV-2A. Ce virus s'apparente à 20 priori, d'après l'analyse des séquences obtenues, à un rétrovirus, mais, étant donnée la technique utilisée pour obtenir cette séquence, il peut aussi bien s'agir d'un virus à ARN dont le génome code pour une enzyme qui possède accessoirement une activité transcriptase inverse 25 comme c'est le cas, par exemple, pour le virus de l'hépatite B, HBV (Shih et coll. J. Virol. 1989; 63, 64-75).

EXEMPLE 6: OBTENTION DE CLONES, DEFINISSANT UNE  
FAMILLE MSRV-2B, PAR AMPLIFICATION PCR "NICHEE" DES  
30 REGIONS POL CONSERVÉES DES RETROVIRUS SUR DES PRÉPARATIONS  
DE VIRIONS ISSUES DES LIGNEES LM7 ET PLI-2.

Une technique PCR dérivée de la technique publiée par Shih et coll. a été utilisée. Cette technique permet, par traitement de tous les composants du milieu 35 réactionnel par la DNase, d'éliminer toute trace d'ADN contaminant. Elle permet parallèlement, par l'utilisation

d'amorces différentes mais chevauchantes dans deux séries successives de cycles d'amplifications PCR, d'augmenter les chances d'amplifier un ADNc résultant d'une quantité d'ARN faible au départ et encore réduite dans l'échantillon par l'action parasite de la DNase sur l'ARN; ce, d'autant plus la DNase est utilisée dans des conditions d'activité en excès qui permettent d'éliminer toute trace d'ADN contaminant avant inactivation de cette enzyme restant dans l'échantillon par chauffage à 85°C pendant 10 minutes. Cette variante de la technique PCR décrite par Shih et coll, a été utilisée sur des fractions de virions, purifiés, comme décrit dans l'exemple 2, comme précédemment pour les surnageants congelés des cultures LM7, issus cette fois de lignée lymphoblastoïde B spontanée provenant d'un nouveau cas de SEP, de la culture PLI-2 (ECACC n° 92072201) et de la culture LM7PC (ECACC n°93010817), ces deux dernières cultures ayant été obtenues selon les méthodes ayant fait l'objet des brevets n° WO 93/20188.

Après clonage avec le TA cloning kit® des produits amplifiés par cette technique et analyse de la séquence à l'aide du séquenceur automatique selon ce qui est décrit dans les exemples 1 et 2, les séquences ont été analysées à l'aide du logiciel Geneworks®, sur la dernière version (79, novembre 93) à ce jour de la banque de données Genebank®.

Cette technique a d'abord été appliquée à deux fractions de virions purifiés sur saccharose à partir de l'isolat "POL-2" produit par la lignée PLI-2 d'une part, et à partir de l'isolat MS7PG produit par la lignée LM7PC d'autre part. Les séquences clonées et séquencées à partir de ces deux échantillons correspondent à quatre catégories. Une première catégorie de séquence, correspond à des séquences amplifiées à partir de matériel nucléique rétroviral contaminant la transcriptase inverse du MoMuLV utilisée pour l'étape de synthèse de l'ADN complémentaire,

ainsi que cela a déjà été signalé précédemment. La deuxième catégorie, retrouvée dans la majorité des clones (55% des clones issus des isolats POL-2 des culture PLI-2, et, 67% des clones issus des isolats MS7PG des cultures 5 LM7PC) correspond à une famille de séquences "pol", dénommées par la demanderesse MSRV-1B et décrites dans les exemples 3 et 4, proches mais différentes du rétrovirus endogène humain dénommé ERV-9 ou HSERV-9, cependant dans une autre région que l'homologie mentionnée précédemment 10 avec la séquence du clone PSJ17 définissant la famille MSRV-1A comme décrit dans les exemples 1 et 2. La troisième catégorie correspond à des séquences très fortement homologues à la séquence attribuée à un virus nommé par la demanderesse MSRV-2A et obtenue précédemment à partir des surnageants 15 de culture LM7 (voir exemple 5) par la technique non modifiée de shih et coll. Cependant la proportion des clones de cette catégorie est inférieure à celle qui a été trouvée précédemment sur l'isolat viral issu de la culture LM7 (5% des clones issus des isolats POL-2 des culture 20 PLI-2, et, 6% des clones issus des isolats MS7PG des cultures LM7PC). La quatrième catégorie correspond à diverses séquences trouvées généralement en un exemplaire, ainsi qu'une de type rétroviral endogène, cependant jamais retrouvées par d'autres approches ou dans d'autres 25 échantillons provenant de culture exprimant une activité transcriptase inverse de type LM7.

La troisième catégorie de séquences obtenues dans cette expérience, est identique au séquence "pol" de type MSRV-2 obtenues précédemment, à quelques mutations près 30 qui peuvent s'expliquer par une variabilité normale du génome viral et par des erreurs des enzymes transcriptase inverse et taq-polymérase utilisées dans cette technique. Cependant, lors de cette deuxième approche utilisant une modification de la technique décrite par Shih et coll., 35 une duplication a priori artéfactuelle de l'amorce aval ("amorce 3') a été observée. Ceci peut provenir de la

configuration des amorces chevauchantes utilisées pour les deux séries de cycles d'amplification PCR (technique "semi-nested). La séquence consensus de ces clones est présentée dans la figure 15. La trame de lecture fonctionnelle correspondante, une représentation de la duplication de l'amorce aval (amorce 3') ainsi que la séquence acides aminés visualisant le motif interne aux régions conservées dans la plupart des rétrovirus (VLPQG et YVDD) sont présentées dans la figure 16. Une comparaison des séquences acides aminés de type MSRV-2B obtenues lors de cette expérience et lors de la précédente fait apparaître une seule différence dans la région communément clonée, à savoir le remplacement de la méthionine M retrouvée dans la région de l'amorce PCR par une valine V, substituant ainsi le motif YVDD, au motif YMDD.

L'analyse dans les banques de données de cette dernière séquence, obtenue par cette technique modifiée sur des échantillons viraux distincts, ne modifie en rien le fait qu'aucune homologie significative n'est trouvée avec des séquences déjà répertoriées dans la base de données Genebank®. Cependant, une analogie certaine de la séquence acides aminés avec la région équivalente des rétrovirus connus existe, ainsi que cela est présenté dans la figure 17. Cette analogie fait intervenir, comme dans les clones issus de la technique précédente, une délétion suivant l'amorce amont. Cette séquence est, ici aussi, compatible avec un génome rétroviral inconnu ou avec un génome viral ARN (transcrit en ADNc dans notre protocole) codant pour une enzyme ayant une activité transcriptase inverse. En effet, les virus de l'hépatite B et de la mosaïque du chou-fleur (CaMV) qui ne sont pas des rétrovirus mais qui possèdent une polymérase à activité transcriptase inverse ont pu être amplifiés selon la technique de Shih et coll. De plus, comme cela est aussi montré dans la figure 17, le CaMV (clone CaMV QG-YM) possède une délétion dans cette région mais, contrairement

à la séquence de type MSRV-2, elle se situe du côté de l'amorce aval (amorce 3').

**EXEMPLE 7: OBTENTION DE CLONES, DEFINISSANT UNE  
5 FAMILLE MSRV-2B, PAR AMPLIFICATION PCR "NICHEE" DES  
REGIONS POL CONSERVÉES DES RETROVIRUS SUR DES PRÉPARATIONS  
DE LYMPHOCYTES B D'UN NOUVEAU CAS DE SEP.**

Enfin, la même technique PCR modifiée d'après la 10 technique de Shih et coll. a été utilisée pour amplifier et séquencer le matériel nucléique ARN présent dans une fraction de virions purifiés au pic d'activité transcriptase inverse "de type LM7" sur gradient de saccharose, selon les protocoles décrits dans les exemples 15 1 et 2, à partir d'une lignée lymphoblastoïde spontannée, obtenue par auto-immortalisation en culture de lymphocytes B d'un patient SEP séropositif pour le virus d'Epstein-Barr (EBV) après mise en culture des cellules lymphoïdes sanguine dans un milieu de culture approprié contenant une 20 concentration appropriée de cyclosporine A. Une représentation de l'activité transcriptase inverse dans les fractions de saccharose prélevées sur un gradient de purification des virions produits par cette lignée (Duc.) est présentée dans la figure 10. De même, les surnageants 25 de culture d'une lignée B obtenue dans les mêmes conditions à partir d'un témoin exempt de sclérose en plaques ont été traités dans les mêmes conditions et le dosage de l'activité transcriptase inverse dans les fractions du gradient de saccharose s'est avéré négatif 30 partout (bruit de fond) et est présenté dans la figure 11. La fraction 3 du gradient correspondant à la lignée B de SEP et la même fraction sans activité transcriptase inverse du gradient témoin non-SEP, ont été analysées par la même technique RT-PCR que précédemment, dérivée de Shih 35 et coll., suivie des mêmes étapes de clonage et de séquençage, comme décrit dans les exemples 1 et 2.

L'analyse des clones recombinants prélevés au hasard a fourni les mêmes quatres catégories de séquences que décrites dans les exemples 3,4 et 6. Les résultats sont présentés dans le tableau 2. Il est tout à fait 5 notable que les séquences de type MSRV-2 soient retrouvées dans le seul matériel associé à un pic d'activité transcriptase inverse "de type LM7" provenant de la lignée lymphoblastoïde B de SEP. Ces séquences n'ont pas été retrouvées avec le matériel de la lignée 10 lymphoblastoïde B témoin (non-SEP) dans 26 clones recombinants pris au hasard. Seules les séquences contaminantes de type Mo-MuLV et des séquences sans analogie rétrovirale particulière ont été retrouvées chez ce témoin. La différence de résultats est à l'évidence 15 hautement significative (chi-2, p<0,001). L'analyse des séquences de type MSRV-2B obtenues à partir de la lignée B de ce patient SEP est présentée dans la figure 18. L'analyse des séquences de type MSRV-2 obtenues dans toutes les fractions de virions purifiées provenant 20 respectivement des cultures LM7, LM7PC, PLI-2 et de la lignée lymphoblastoïde B de SEP (Duc.), par les deux techniques PCR dérivée de la technique publiée par Shih et coll., est présentée dans la figure 19. On y constate que la séquence en acide nucléique y est très conservée, à 25 quelques mutations près, et que la seule variation de séquence acides aminés observée dans ces clones concerne la méthionine ou la valine du site amorce 3' YMDD/YVDD.

EXAMPLE 8: Amplification PCR des la séquence 30 nucléique contenue entre la région 5' définie par le clone "pol MSRV-1B" et la région 3' définie par le clone PSJ17 (MSRV-1A).

Cinq oligonucléotides, M001, M002-A, M003-BCD, 35 P004 et P005, ont été définis pour amplifier de l'ARN provenant de virions purifiés POL-2. Des réactions de

contrôle ont été effectuées de manière à contrôler la présence de contaminants (réaction sur de l'eau). L'amplification consiste en une étape de RT-PCR selon le protocole décrit dans l'exemple 2 suivie d'une PCR 5 "nichée" selon le protocole PCR décrit dans l'exemple 2. Dans le premier cycle RT-PCR, les amorces M001 et P004 ou P005 sont utilisées. Dans le deuxième cycle PCR, les amorces M002-A ou M003-BCD, et l'amorce P004 sont utilisées. Les amorces sont positionnées comme suit:

10

	M002-A				
	M003-BCD				
M001	—		P004	P005	
—	—		—	—	
15	<hr/>				ARN
POL-2	<----->				
<----->	<----->				
pol MSRV-1					SJ17

20

Leur composition est:

amorce M001: GGTCITICCICAIGG

amorce M002-A: TTAGGGATAGCCCTCATCTCT

amorce M003-BCD: TCAGGGATAGCCCCATCTAT

amorce P004: AACCCTTGCCACTACATCAATTT

25

amorce P005: GCGTAAGGACTCCTAGAGCTATT

L'alignement du consensus MSRV-1B avec le produit d'amplification "nichée" décrit ci-dessus et nommé M003-P004 et le clone PSJ17 (MSRV-1A) montrant la communauté de 30 famille des isolats MSRV-1A et MSRV-1B, ainsi que la séquence nucléique de la région comprise entre MSRV-1B et MSRV-1A sont présentés dans la figure 20.

Il convient ici de rappeler que, dans le modèle naturel animal de la SEP représenté par le Visna du 35 mouton, la co-existence, le co-isolement, la co-purification et même, dans certaines situations

pathologiques étudiées sous cet angle, la coopération, de deux virus, les rétrovirus Maedi-Visna et JSRV (réetrovirus de l'adénomatose ovine pulmonaire) est connue et fréquemment mise en évidence dans les maladies naturelles 5 (et non-expérimentales), ainsi que l'ont rapporté De Martini J.C. et coll. (J.N.C.I. 1987; 79, 167-177), Rosadio R.H. et coll. (Vet. Pathol. 1988; 25, 58-66), Dawson M. et coll. (Br. Vet. J. 1990; 146, 531-537) et York D.F. et coll. (J. Virol. 1991; 65, 5061-5067).

10 Or, le virus Maedi-Visna est un rétrovirus exogène directement responsable de la pathologie neurologique du mouton alors que le JSRV est un rétrovirus endogène responsable d'un cancer, l'adénomatose ovine pulmonaire. La coopération entre ces deux virus a été étudiée et 15 montrée dans une troisième pathologie du mouton, le maedi qui est une forme de pneumonie interstitielle lymphocytaire, mais cette synergie n'a pas été étudiée ni recherchée dans le cas du Visna, la maladie démyélinisante du système nerveux central.

20 Dans ces conditions, on peut considérer que la présence de deux éléments viraux différents, co-isolés dans une pathologie naturelle humaine telle que la SEP, trouve un précédent logique du point de vue conceptuel dans le modèle du mouton évoqué ci-dessus.

25 Ces virus peuvent donc être à l'origine de pathologies différentes mais aussi coopérer dans certains processus pathogéniques.

Dans ces conditions on doit concevoir que les applications diagnostiques et thérapeutiques découlant de 30 la découverte des virus MSRV-1 et MSRV-2 contenant les séquences définies dans le présent brevet, aux variations près que ces virus peuvent naturellement présenter, pourraient se recouper dans le cadre de la SEP.

Par exemple, des rétrovirus endogènes de type 35 MSRV-1 pourrait interférer de façon variable, selon les génotypes MSRV-1 portés par des individus différents, avec

l'infection par un virus exogène de type MSRV-2 qui serait plus spécifiquement associé à l'induction de la maladie sclérose en plaques. Ils pourraient donc moduler la pathogénicité de ce dernier. Dans ces conditions,

5      génotypage des gènes de type MSRV-1 portés par un individu dans une famille ou une région à risque, permettrait de faire un diagnostic de "risque individuel". Le diagnostic d'une infection à MSRV-2, devrait être confronté à ce typage MSRV-1, pour évaluer le risque de maladie clinique.

10     De plus, certaines formes graves ou, à contrario bénignes, de SEP pourraient être caractérisées par la présence d'un haplotype MSRV-1 endogène particulier, et ainsi, un pronostic sur l'évolution clinique pourrait être donné. De plus, de tels éléments génétiques pourraient faire l'objet

15     d'une thérapeutique complémentaire utilisant les procédés de la thérapie génique ou, plus simplement, des inhibiteurs de l'expression rétrovirale actifs sur des rétrovirus endogènes.

Selon un autre exemple, l'expression d'un rétrovirus endogène de type MSRV-1 pourrait être transactivée par une protéine régulatrice codée et exprimée par un virus de type MSRV-2. Ainsi, on peut concevoir dans ces conditions qu'un diagnostic devrait dépister une infection par MSRV-2 mais aussi une réaction immunitaire particulière contre (ou une expression particulière de) MSRV-1, pour être à même de caractériser l'état de "sclérose en plaques active". Ainsi encore, une réponse immunitaire contre MSRV-2 ou la détection directe d'un de ses antigènes ou d'une séquence nucléique lui appartenant, en l'absence d'expression anormale d'un endogène de type MSRV-1, pourrait caractériser un stade "latent" de la maladie. Aussi, le contrôle régulier direct (déttection ARN ou protéines virales) ou indirect (immunité) de l'état d'expression de populations endogènes MSRV-1, pourrait constituer un contrôle pronostic extrêmement précoce de passage à la maladie clinique si

celle-ci s'accompagnait d'une modification caractéristique de cette expression. Ce type de suivi "diagnostic ou pronostic" pourrait permettre de rationaliser une approche thérapeutique et d'en optimiser l'efficacité.

5 Selon un autre exemple, ces deux virus pourraient être exogènes tous les deux, bien que l'un d'entre eux soit phylogénétiquement proche d'une famille de rétrovirus endogène, et présenter une synergie ainsi que cela peut se rencontrer notamment dans le cas bien connu de l'hépatite 10 delta, dont l'agent causal, le virus VHD possède un génome défectif dépendant du virus VHB, agent de l'hépatite B. Cet exemple connu démontre que deux maladies affectant le même organe et différenciables, pour l'étude clinique, essentiellement sur des critères de gravité, peuvent 15 coexister par superposition de l'une (hépatite delta) sur l'autre (hépatite B) dont la présence est requise du fait de la dépendance du génome du VHD par rapport au VHB. Un tel cas de figure, ou tout autre type d'interaction, peut donc se concevoir dans le cas de deux virus exogènes, et 20 même d'un rétrovirus endogène, associés à la SEP. Dans un cas similaire à celui des hépatites B et delta, on pourrait dissocier, par les diagnostics respectifs liés à MSRV-1 et MSRV-2, des formes de SEP avec un seul virus associé, et des formes, à priori plus graves, avec deux 25 virus associés.

Selon un autre exemple qui s'inspire d'un raisonnement analogique étayé sur la connaissance des hépatites virales, il est aussi concevable que plusieurs virus indépendants puissent provoquer un syndrome 30 inflammatoire du système nerveux central, dont les variantes cliniques sont regroupées sous la dénomination "sclérose en plaques", ainsi qu'il existe des "hépatites" causées alternativement par le VHA, le VHB, le VHC, le VHE ou l'association VHB et VHD. Dans ces conditions, le 35 diagnostic des différents virus associés à une "sclérose en plaques" permettrait de préciser le type de virus en

cause et, par là, les mesures de prophylaxie et les méthodes thérapeutiques à utiliser pour chaque malade, alors qu'un traitement anti-inflammatoire ou immunodépresseur global est actuellement proposé sans 5 discrimination efficace des causalités de la maladie, chez l'ensemble des malades atteints de SEP.

Ces virus peuvent aussi être à l'origine d'autres pathologies individualisées, et notamment, dans le cas de MSRV-1 dont l'expression est nettement accrue dans les 10 cellules immortalisées ou transformées des lignées continues par rapport aux mêmes cellules en culture primaire, une association avec des processus tumoraux, certains cancers ou syndromes hyperprolifératifs, est concevable.

15 Alternativement, un des deux virus peut s'avérer, seul et à l'exclusion de l'autre, jouer un rôle dans la pathogénie de la SEP et, donc, seules les applications diagnostiques et thérapeutique de celui-ci seraient d'intérêt pour la sclérose en plaques. Mais, comme dans 20 l'exemple du visna et du jaagsiekte précité, il se pourrait que l'autre virus puisse indépendamment induire un tout autre type de maladie chez l'homme.

TABLEAU 1

## Principaux rétrovirus endogènes humains

Nom	Longueur (kb)	Nombre de copies	Relation avec les rétrovirus exogènes	Références
4.1	8.8	5-100	MoMuLV	Martin, PNAS, 1981
51.1	6	35-50	MoMuLV	Science, 1984
ERV.1	3-4	1	MoMuLV/BaEV	Bonner, PNAS, 1982
ERV.3	9.9	1	BaEV	O'Connell, Virology, 1984
MM.2	9	50	MMTV	Callahan, PNAS, 1982
HERV.K	9.5	50	MMTV	Onc, J. Virol., 1987
HERRS.P	8-9	20-40	MoMuLV	Kroger, J. Virol., 1987
RTLV-II	5.8	1000	HTLV-1/BLV	Maier, J. Virol., 1987
HERS.I	5-6	1	HTLV-1	Perl, Genomics, 1991
M7-ret	4-10	300	M1 baboon RV	Noda, Nucl. Acids Res., 1982
S71	6	1	SSAV	Werner, Virology, 1990
ERV.9	8	35-40	?	I.a Mantia, Nucl. Acids Res., 1991
Hs5	5-10	?	FelV/Fv.4	Levy, J. Gen. Viro., 1990
EIS.1	3.2-9.5	?	HTLV-1	Horwitz, J. Virol., 1992
EIS.2	3.2-9.5	?	HTLV-1	

TABLEAU 2

REPARTITION DES SEQUENCES OBTENUES PAR PCR  
MODIFIEE D'APRES SHIH ET COLL. A PARTIR DE CULTURE  
DE LYMPHOCYTES B D'UN PATIENT ATTEINT DE SCLEROSE  
EN PLAQUES VERSUS UN PATIENT CONTROLE

	TYPE 1	TYPE 2	TYPE 3 MSRV-2B	TYPE 4
Lympho B patient SEP	0 0%	11 48%	4 17%	8 35%
Lympho B patient non SEP	5 19%	0 0%	0 0%	21 81%

## REVENDICATIONS

1/ Association de deux virus associés à la sclérose en plaques, à savoir, un premier virus humain, possédant une activité transcriptase inverse, et associé à une famille d'éléments rétroviraux endogènes, ou un variant du premier virus comprenant au moins un antigène reconnu par au moins un anticorps dirigé contre au moins un antigène correspondant dudit premier virus, et un second virus, possédant une activité transcriptase inverse, ou un variant dudit second virus, ces deux virus étant issus ensemble d'une même souche virale choisie parmi les souches dénommées respectivement POL-2, déposée le 22.07.1992 auprès de l'ECACC sous le numéro d'accès V92072202 et MS7PG déposée le 08.01.93 auprès de l'ECACC sous le numéro d'accès V93010816, et parmi les souches variantes consistant en des virus comprenant au moins un antigène qui est reconnu par au moins un anticorps dirigé contre au moins un antigène correspondant de l'un ou l'autre des virus des souches virales POL-2 et MS7PG précitées.

2/ Association de deux virus associés à la sclérose en plaques, à savoir, un premier virus humain possédant une activité transcriptase inverse, et associé à une famille d'éléments rétroviraux endogènes, ou un variant du premier virus comprenant au moins un antigène reconnu par au moins un anticorps dirigé contre au moins un antigène correspondant dudit premier virus, et un second virus possédant une activité transcriptase inverse, ou un variant du second virus, ces deux virus étant produit par une même lignée cellulaire choisie parmi les souches dénommées respectivement PLI-2 déposée le 22.07.1992 auprès de l'ECACC sous le numéro d'accès 92072201 et LM7PC déposée le 08.01.93 auprès de l'ECACC sous le numéro d'accès 93010817, et par toutes cultures cellulaires infectées susceptibles de produire un virus comprenant au moins un antigène qui est reconnu par au

moins un anticorps dirigé contre au moins un antigène correspondant de l'un ou l'autre des virus produits par les lignées PLI-2 et LM7PC précitées.

3/ Association de deux virus, à savoir un premier virus dont le génome comprend une séquence nucléotidique présentant au moins 50 % et préférentiellement au moins 70 % d'homologie avec la séquence nucléotidique SEQ ID N01 ou sa séquence complémentaire, ou un variant dudit premier virus dont le génome code pour une séquence peptidique présentant au moins 50 % et préférentiellement au moins 70 % d'homologie avec la séquence peptidique codée par la séquence nucléotidique SEQ ID N05 ou sa séquence complémentaire, et un second virus, dont le génome comprend une séquence nucléotidique présentant au moins 50 % et préférentiellement au moins 70 % d'homologie avec une séquence nucléotidique choisie parmi SEQ ID N02 ou sa séquence complémentaire, ou un variant dudit second virus dont le génome code pour une séquence peptidique présentant au moins 50 %, et préférentiellement au moins 70 % d'homologie avec la séquence peptidique codée par la séquence nucléotidique SEQ ID N02 ou sa séquence complémentaire

4/ Procédé de détection d'un premier virus, et/ou d'un second virus, associés à la sclérose en plaques, caractérisé en ce qu'on met en oeuvre au moins deux fragments, un premier fragment dont la séquence nucléotidique présente au moins 50 % et de préférence au moins 70 % d'homologie avec SEQ ID N05 ou sa séquence complémentaire, et/ou un second fragment dont la séquence nucléotidique présente au moins 50 % et de préférence au moins 70 % d'homologie avec SEQ ID N02 ou sa séquence complémentaire, chacun desdits fragments étant une sonde ou une amorce.

5/ Composition diagnostique, prophylactique ou thérapeutique, caractérisée en ce qu'elle comprend une association selon la revendication 4 et un premier

fragment nucléotidique, et/ou un second fragment nucléotidique, tels que définis selon la revendication 4.

6/ Procédé pour détecter et/ou identifier une association de virus associés à la sclérose en plaques, 5 dans un échantillon biologique, caractérisé en ce qu'on expose un ARN et/ou un ADN spécifique à chaque dit virus, et/ou leur ARN et/ou ADN complémentaire à une association d'un fragment nucléotidique et d'un second fragment nucléotidique, tels que définis à la revendication 4.

10 7/ Procédé de détection d'un premier virus, et/ou d'un second virus, associés à la sclérose en plaques, caractérisé en qu'on met en oeuvre un premier peptide codé de manière partielle ou totale par le premier fragment nucléotidique défini à la revendication 4, et/ou un second 15 peptide codé de manière partielle ou totale par le deuxième fragment nucléotidique défini à la revendication 4.

8/ Composition diagnostique, prophylactique ou thérapeutique, caractérisée en ce qu'elle comprend un 20 premier peptide et/ou un second peptide, définis à la revendication 7.

9/ Composition diagnostique, prophylactique ou thérapeutique, caractérisée en ce qu'elle est obtenue par préparation d'un premier ligand spécifique du premier 25 peptide, et/ou un second ligand spécifique du second peptide, et par association du premier ligand et/ou du second ligand.

10/ Fragment nucléotidique, caractérisé en ce que sa séquence nucléotidique présente au moins 50%, et de 30 préférence au moins 70% d'homologie, avec une séquence choisie parmi SEQ ID N01, SEQ ID N03, SEQ ID N04, et la séquence SEQ ID N05 comprenant bout à bout SEQ ID N01, SEQ ID N03, SEQ ID N04.

## FIG 1

CAAGCCACCC AAGAACCTTT AAATTTCCTC ACTACCTGIG GCTACAAGGT	50
TTCCAAACCA AAGGCTCAGC TCTGCTCACA GGAGATTAGA TACTTAGGTT	100
TAAAATTATC CAAAGGCACC AGGGCCTCA GTGAGGAACG TATCCAGCCT	150
ATACTGGGTT ATCCTCATCC CAAAACCTA AAGCAACTAA GAGGGTTCCT	200
TAGCATGATC AGGTTCTGCG CGAAAACAAG ATTCCCAGGT ACAACCAAAA	250
TAGCCAGACC ATTATATACA CTAATTAAGG AAACCTAGAA AGCCAATACC	300
TATTTAGTAA GATGGACACC TAAACAGAAG GCTTTCAGG CCCTAAAGAA	350
GGCCCTAACCC CAAGCCCCAG TGTTCAGCTT GCCAACAGGG CAAGATTTTT	400
CTTTATATGG CACAGAAAAA ACAGGAATCG CTCTAGGAGT CCTTACACAG	450
GTCGGAGGGA TGAGCTTGCA ACCCGTGGCA TACCTGAATA AGGAAATTGA	500
TGTAGTGGCA AAGGGTTGGC CTCATNGTT ATGGGTAAATG GNGGCAGTAG	550
CAGTCINAGT ATCTGAACCA GTTAAAATAA TACAGGGAAAG AGATCTINCT	600
GIGTGGACAT CTCATGATGT GAACGGCATA CTCACTGCTA AAGGAGACTT	650
GIGGTGTCGA GACAACCATT TACTTAANTA TCAGGCTCTA TIACTTGAAG	700
AGCCAGTGCT GNGACTGCGC ACTTGTGCAA CTCTAAACC C	741

## FIG 2

Consensus	CACGCCACYC WAGAWCTCTT RAAAYTTYCTIM RCTAMYRWC GSTACAAAGGT	50
PSJ17	..A..C..C. A...A..... A..T..C..C A...CCIGT. .C.....	50
ERV9 pol	..G..T..T. T...T..... G..C..T..A G...ATCAA. .G.....	50
Consensus	KTCYAWRYCR AAGGCGYCAGC TYTGCYYACA CGAGATTAGA TACTYAGGT	100
PSJ17	T..C..AAC..A .....T.... C...TC... .....T.....	100
ERV9 pol	G..T..TGT..G .....C.... T...CT... .....C.....	86
Consensus	TAAATATCTA GGCTTAATCT TAKCCAAAGG SACCAGGGSC CTCAGYRAGG	150
PSJ17	.....T..... C.....G.....TG...	136
ERV9 pol	.....G.....G.....C.....CA...	136
Consensus	AAYGNAATMCA GCCTATACTG GSTTATCCTC RYCCYAARAC MYTAAARCAR	200
PSJ17	..C..T..C.. .....G..... AT..C..A.. CC.....G..A	186
ERV9 pol	..T..A..A..... C..... GC..T..G.. AT.....A..G	186
Consensus	YTRAGQRGRT TCTTTRGMAT KAYCAGGYTT YTCCCGAMWA YARGATYCCY	250
PSJ17	C.A..A.G.. ....A.C.. G.T...T.. C.....AA. C.A...T..C	236
ERV9 pol	T.G..G.A.. ....G.A.. T.C..C.. T.....CT. T-G...C..T	234
Consensus	RGRTACAASC RARAYAGCCA GRCGMYTATA TACACTAATY AAGGAAACYC	300
PSJ17	A.G.....C. A.A.T..... A..AT..... T.....T.	286
ERV9 pol	G.A....G. G.G.C..... G..CC...C.....C.....C.	276
Consensus	AGARRCMAA TACYATTA GTTRAGWKGG AMMCMKARRC AGAAACRRCY	350
PSJ17	...AA..C...CT..T.. ..A..AT.. .CA..CT..AA.. ....GG..T	333
ERV9 pol	...CG..A...TC..C.. ...G..TG.. .AC..AG..GG.. ....AAC	326
Consensus	TTCMARRCCY TAAAGMAGGC YCTARYMCAA GCYCCAGYKT TMCAGCYIKCC	400
PSJ17	...C..GG..C ....A.... C...ACC... ..C....TG. .C...T.G..	383
ERV9 pol	...A..AA..T ....C.... T...GTA... ..T....CT. .A...C.T..	376
Consensus	MACAGGRCAR RATTTCCTT TATATGKCAAG AGARARARCM RGRATMGCTC	450
PSJ17	A.....G..A.....T.....G..C ..A.A.A.A G.A..C....	433
ERV9 pol	C.....A..G A....C.....T..G ...G.G.G.C A.G..A....	426
Consensus	TWGGAGTCTT YACWCARRYY CRWGGGAYRA SCYVRCAAAC MGTTGGCATAAC	500
PSJ17	..A..... T..A..GGTC ..CA... TG. G..TTG..... C.....	483
ERV9 pol	.T..... C..T..AACT ..AT.....CA. C.CCA..... A.....	476
Consensus	CTRARTAAGG AAATTCATGT AGTRGCAAAR CGYTGCGCTC AYNGTTTAWG	550
PSJ17	..G.A..... .....G.....G ..T..... T.....T.	533
ERV9 pol	..A.G..... .....A.....A ..C..... .CT.....A.	526
Consensus	GGTARTKGNR GCAGTRGCGC TCTINAGYATC WGARGCWRTY AAAATAATAC	600
PSJ17	....A.G..G .....A..A.....T... T..A..AG..T ..	583
ERV9 pol	....G.T..CA .....G..C.. ...T..C.. A..G..TA..C ..	576
Consensus	ARGGAARRGA TCTYNCTGTS TRGACWMCY ATGATGTRAA YGCCATACTM	650
PSJ17	..G....GA.. ...T.....G. G...AT..C .....G.. C.....C	633
ERV9 pol	.A....AG.. ...CA..C A..TA..T .....A.. T.....A	625
Consensus	RSTGCYAAAG GARRYTIRIG GYTRTCAGAC AACYRYYTAC TIAARNYAVCA	700
PSJ17	AC...T..... GAC..G.. .T.G..... .CATT... ..A..T..	683
ERV9 pol	GG...C..... AGT..A.. .C.A..... .TCCC... ..GA..C..	675
Consensus	GGCWCTAYIM CTTGARGRRC CAGTOCTKNR AMIRYGCACW TGYRYRRCCYC	750
PSJ17	...T...T.A ..A..AG. ....G.G ..C..GC....T ..TCCAA..T.	733
ERV9 pol	...A...C..C ..G..GA..... TCA ..A..AT....A ..CATGG..C.	725
Consensus	TYAAMCCY 758	
PSJ17	.T..A..C 741	
ERV9 pol	.C..C..T 733	

2715937

2715937

FIG 3

QATQELLNFL TTICGYKVSKP KAQLCSQEIR YLCIKLSKGT RGLSEERIQP	50
ILGYPHPKTL KQLRGFLSMI RFLPKTRFPG TTIKIAARPLYT LIKETQKANT	100
YLVRWTPKQK AFQALKKALT QAPVFSLPTG QDFSLYGTEK TGIALGVLTQ	150
VRQMSLQPVA YLNKEIDVA KGWPDXLWVM XAVAVXVSEA VKTIQGRDLX	200
VWTSHDVNGI LTAKGDLWLS DHILLXYQAL LLEEPVLXLR TCATLKP	247

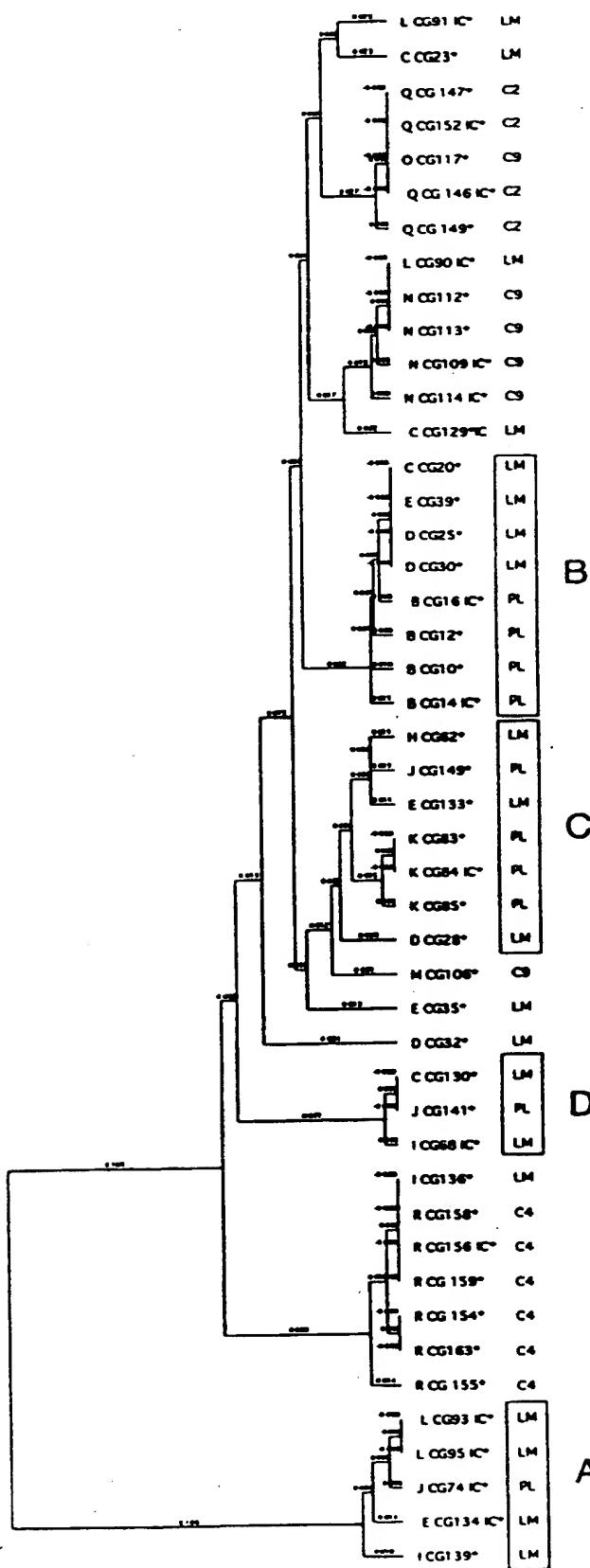
## FIG 4

Consensus	QAT..LLNFL ...GYKVS.. KAQLC.Q... YLGL.L.KGT R.LS.E.IQP	50
Trad PSJ17	...QE..... TTC..... KP ..... S.EIR ....K.S... .G..E.R...	50
Trad ERV9 modif	...LD..... ANQ..... MS ..... L.QVK ....I.A... .A..K.X...	50
Consensus	IL.YP.PKTL KQLR.FL... .F..... PG ....ARP... LIKETQ.ANT	100
Trad PSJ17	..G..H.... ....G..SMI R.LPKTRF.. TTKI...LYT .....K...	100
Trad ERV9 modif	..A..R.... ....E..GIT S.-CRLWI.. YSET...--- .....R...	96
Consensus	.LV.W.P... ..F..LK.AL .QAP.. SLPT GQ.FSLY..E ...IALGVLT	150
Trad PSJ17	Y..R.T.K-Q KA.QA..K.. T...VF.... ..D....GT. KTG.....	149
Trad ERV9 modif	H..E.E.EAE TT.KT..Q.. V...AL.... ..N....VR. RAR.....	146
Consensus	Q..G...QPV AYL.KEIDVV AKGWP.H.L.V ..AVAV..SE A.KIIQG.DL	200
Trad PSJ17	.VR.MSL.... ..N..... ....X.W. MX....XV.. .V.....R..	199
Trad ERV9 modif	.TH.TTP.... ..S..... ....C.R. VA....LA.. .I.....K..	196
Consensus	...T..DVNG IL.AKG.IWL SDN.LL.YQA LL.E.PVL... .TC..L.P	248
Trad PSJ17	XW.SH.... ..T...D... ....H..X... ....E...XL R..AT.K.	247
Trad ERV9 modif	TS-.TY.... ..G...S... ....C..R... ....G...QI C..MA.N.	243

5/20

FIG 5

FIG 6



## FIG 7

## CONSENSUS A

GTTAGGGATAGCCC TCATCTTTGGTCA GGTACTGGCCAAGA TCTAGGCCACTTC 60  
 V . G . P S S L W S G T G P R S R P L L  
 F R D S P H L F G Q V L A Q D L G H F S  
 L G I A L I S L V R Y W P K I . A T S Q  
  
 AGGTCCAGGCACTCT GTTCCTTCAG 85  
 R S R H S V P S  
 G P G T L F L Q  
 V Q A L C S F

## CONSENSUS B

GTTCAGGGATAGCCC CCATCTATTGGCCA GGCACTAGCTCAATA CTTGAGCCAGTTCTC 60  
 V Q G . P P S I W P G T S S I L E P V L  
 F R D S P H L F G Q A L A Q Y L S Q F S  
 S G I A P I Y L A R H . L N T . A S S H  
  
 ATACCTGGACACTCT TGTCCCTCGGT 86  
 I P G H S C P S  
 Y L D T L V L R  
 T W T L L S F G

## CONSENSUS C

GTTCAGGGATAGCCC CCATCTATTGGCCA GGCATTAGCCAAGA CTTGAGTCATTCTC 60  
 V Q G . P P S I W P G I S P R L E S I L  
 F R D S P H L F G Q A L A Q D L S Q F S  
 S G I A P I Y L A R H . P K T . V N S H  
  
 ATACCTGGACACTCT TGTCCCTTCAG 85  
 I P G H S C P S  
 Y L D T L V L Q  
 T W T L L S F

## CONSENSUS D

GTTCAGGGATAGCTC CCATCTATTGGCCT GGCATTAACCCGAGA CTTAAGCCAGTTCTC 60  
 V Q G . L P S I W P G I N P R L K P V L  
 F R D S S H L F G L A L T R D L S Q F S  
 S G I A P I Y L A W H . P E T . A S S H  
  
 ATACGTGGACACTCT TGTCCCTTGG 85  
 I R G H S C P L  
 Y V D T L V L W  
 T W T L L S F

FIG 8A

Consensus	...SP.LFQ.	..A..L...R	...P.....	30
HIV1 QG/YM	WKGS <del>PA</del> IFQ	SM <del>I</del> KILEPFR	KQNPDIVTYQ	30
HIV2 QG/YM	WKGS <del>PA</del> IFQH	TMRQVLEPFR	KANKD <del>V</del> IIQ	30
VISNA QG/YM	WKLSPAVYQF	TMQKILRGWI	E <del>H</del> PMI <del>Q</del> FGI	30
CAEV QG/YM	WKLSPSVYQF	TMQEILEDWI	Q <del>H</del> PEI <del>Q</del> FGI	30
MMTV QG/YM	MKN <del>S</del> PTLCQK	FVDKAILTVR	DRYQDSYTVH	30
HERV <del>K</del> QG/YM	M <del>N</del> SPTICQT	FVGRALQPVR	EKFSDCYI <del>H</del>	30
JSRV QG/YM	M <del>T</del> N <del>S</del> PTLCQK	FVATAIAPVR	QRF <del>P</del> QLYLVH	30
MPMV QG/YM	M <del>A</del> NSPTLCQK	YVATAIHKVR	HAWKQMYI <del>H</del>	30
HTLV1 QG/YM	FKN <del>S</del> PTLFEM	QLAHILQPIR	QAFPQCTI <del>I</del> Q	30
HTLV2 QG/YM	FKN <del>S</del> PTLF <del>E</del> Q	QLA <del>A</del> VLNPMR	KMFPTISTIVQ	30
Trad cons30 B	FRD <del>S</del> PHLF <del>G</del> Q	ALA <del>Q</del> YLSQF-	-SYLD <del>I</del> LVL <del>R</del>	28
Trad cons33/28 C	FRD <del>S</del> PHLF <del>G</del> Q	ALA <del>Q</del> DLSQF-	-SYLD <del>I</del> LVLQ	28
Trad cons24 D	FRDSSHLF <del>G</del> L	AL <del>I</del> RDL <del>S</del> QF-	-SYVD <del>I</del> LVLW	28
ERV9 QG/YM	FRD <del>S</del> PHLF <del>G</del> Q	ALA <del>K</del> DLGHF-	-SSPG <del>I</del> LVLQ	28
Trad cons37 A	FRD <del>S</del> PHLF <del>G</del> Q	VLA <del>Q</del> D <del>L</del> GHF-	-SGPG <del>I</del> LFLQ	28

FIG 8B

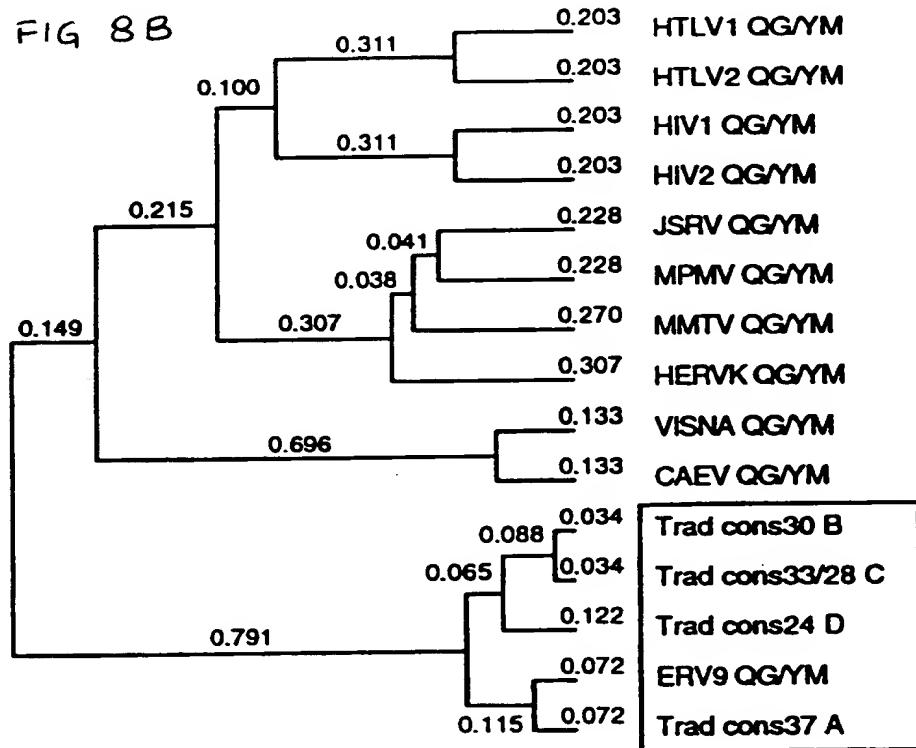


FIG 8C

Consensus	FRDSPHLFGQ	ALAQQLSQFS	Y.DILVILQ	28	
Trad cons30 B	.....	....Y.....	L.....R	28	71%
Trad cons33/28 C	.....	.....	L.....	28	79%
Trad cons24 D	....S....L	..TR.....	V.....W	28	64%
ERV9 QG/YM	.....	...K.GH..	SPG.....	28	100%
Trad cons37 A	.....	V.....GH..	GPG..F..	28	86%

## FIG 9

## CONSENSUS A

Consensus	GTTCAGGGAT ANCCTCATIC TCTTGGTCA GGTACTGCC CAAGATCTAG	50
Consensus	GCCACTTCCTC AGGTOCAGSN ACTCTGTYCC TTCAG	85

## CONSENSUS B

Consensus	GTTCAGGGAT AGCCCCCATC TATTGGCCA GGCACTAGCT CAATACCTGA	50
Consensus	GCCAGTTCTC ATAACCTGGAC AYTCTYGTCC TTGGT	86

## CONSENSUS C

Consensus	GTTCARRGAT AGCCCCCATC TATTGGCCW RGYATTAGCC CAAGACTTGA	50
Consensus	GYCAATTCTC ATAACCTGGAC ACTCTGTCCT TYRG	85

## CONSENSUS D

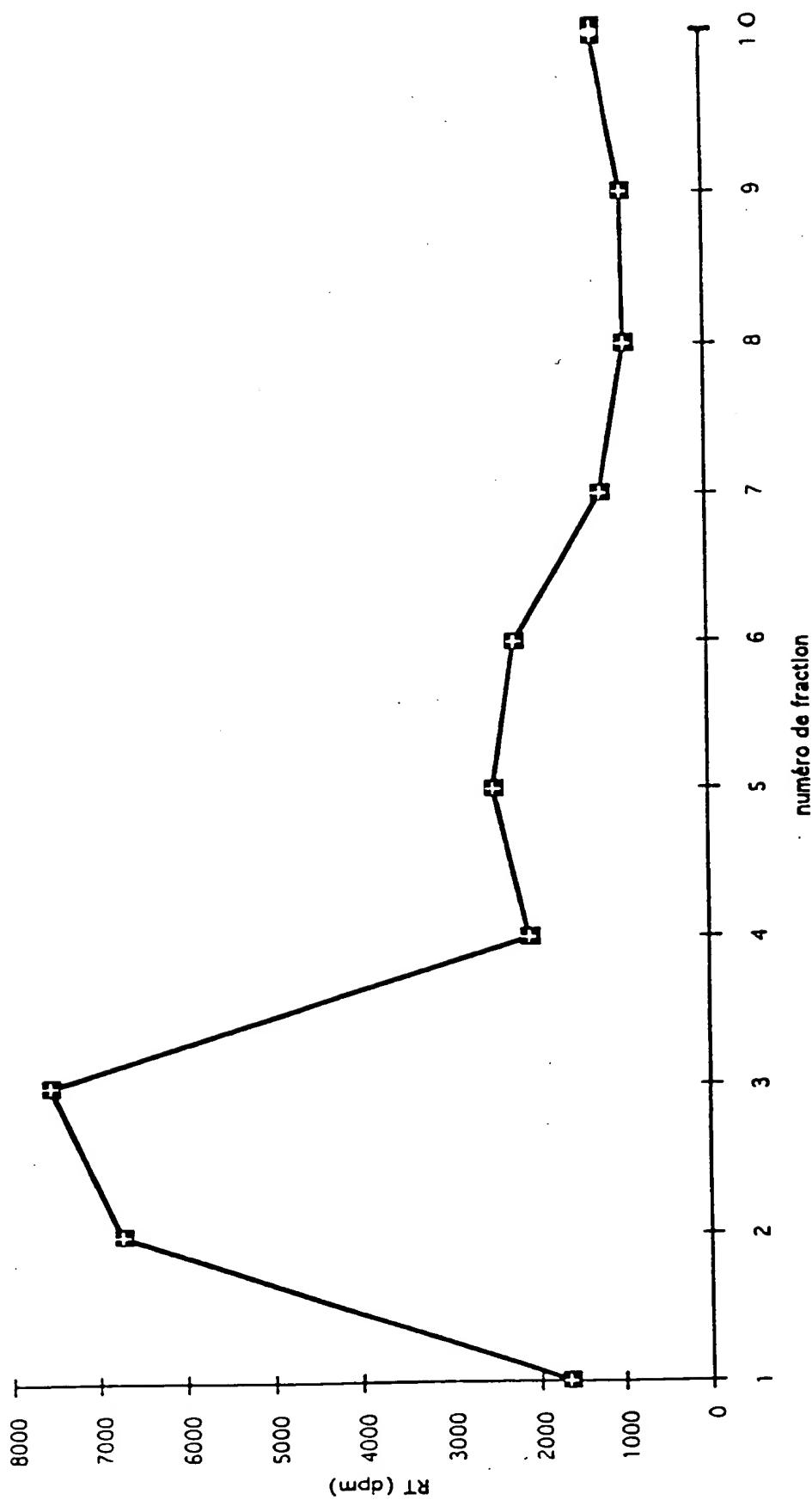
Consensus	GTTCAGGGAT AGCTCCATC TATTGGCTT GGCATTAACC CGAGACTTAA	50
Consensus	GCCAGTTCTY ATAOGTGGAC ACTCTGTCCT TTGG	85

## CONSENSUS GENERAL MSRV-1B AVEC AMORCES D'AMPLIFICATION (SEQ ID NO 3)

Consensus	GIGTIGCCAC AGGGGTTTAR RGATANCYCY CATCIMTTG GYCWRGYAYT
Consensus	RRCYCRAKAY YIRRGYCAVT TCTYAKRVSY RGSNAYCTB KYCCTTYRGT
Consensus	ACATGGATGA C

2715937

FIG 10



11/20

FIG 11

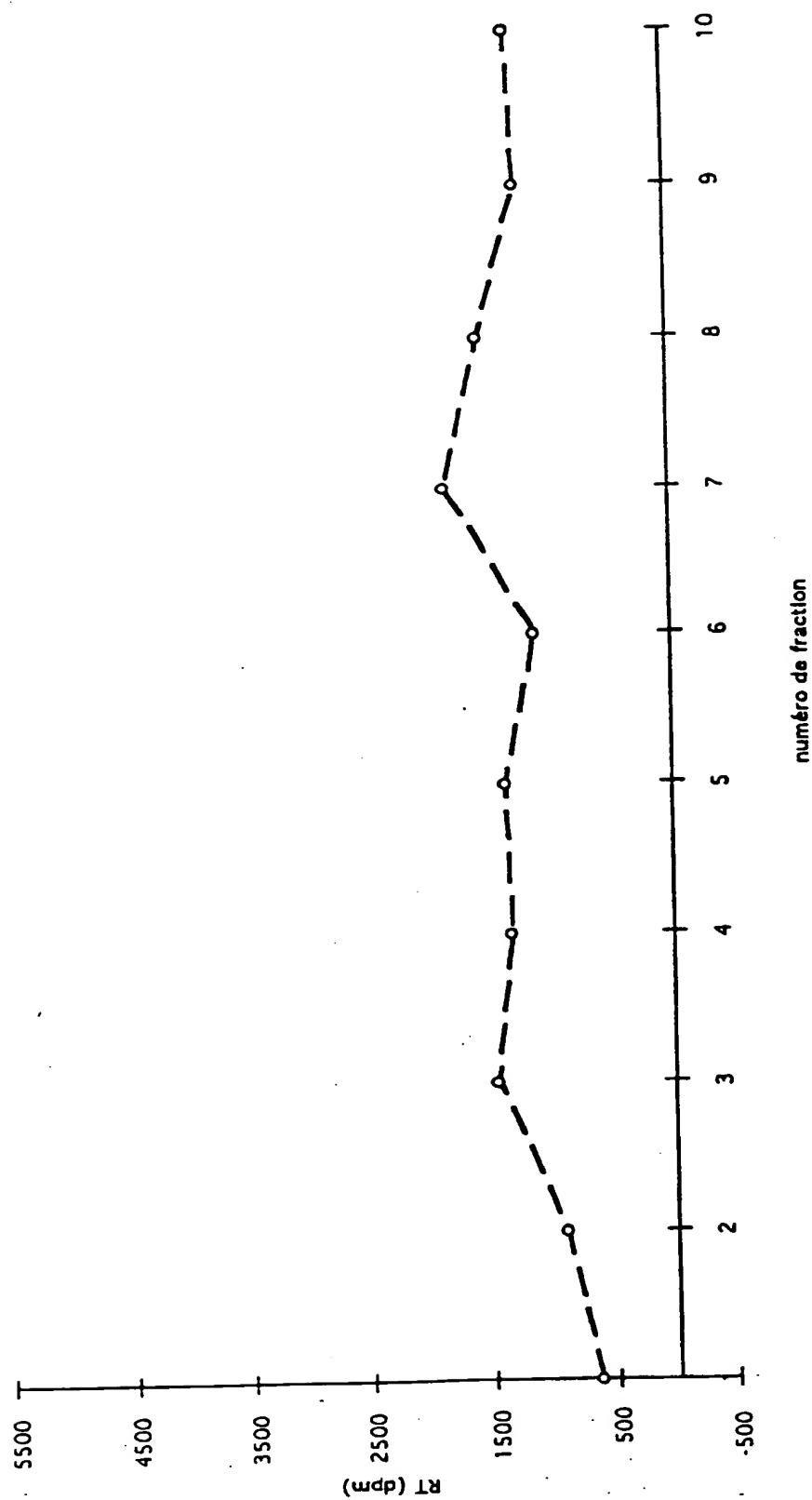


FIG 12A

Consensus	CTTGGATCCA QNGYTVMCAG ARGGGNNCAG GGNRTANNCC CCNICTINTYW	50
MAJ	CTTGGATCCA GtGtTgcCaC AgGGGtTCAG GG-aTAGccC CCaTCTaTtt	50
VAR	.....c.aa.c. .a...c.C.. .gg...a. ....ca	50
Consensus	GGMCAGGSTA TTAGYCCAAG ACTTIGAKCCA GNNCTCATAAC CTGGAACACT	100
MAJ	GGcCAGGc-A TTAGcCCAAG ACTTIGaGCCA GtTCICATAAC CTGGA-CACT	100
VAR	..a....gt. ....t..... ....a....	98
Consensus	CTTGCNCCTT CGGNACGRKG GNATGACMIN CTGANGCTTG AGA	143
MAJ	CTTG-TCCCTT CGGTAC-atG G-ATGACcTg CTGAagCTTG AG-	141
VAR	....c---- ....gggG .c.....a.c CTGANGCTTG AG-	137
MIN	.....----- .....	125

FIG 12B

an	CTTGGATccA <u>GTGTTGCCAC AGGGGTTCAAG GGATAGCCCC CATCTATTG</u>	50
prot	V L P O G F R D S P H L F G	
an	<u>GCCAGGCATT AGCCCCAACAC TTGAGCCAGT TCICATACCT GGACACICIT</u>	100
prot	Q A L A Q D L S Q F S Y L D T L	
an	<u>GTCCTTCGGT ACATGGATGA CCTGCTGaaG CTTGAG</u>	137
prot	V L R Y M D D	

## FIG 13

pol SHIH TGGAAAGTGT TGCCACAGGG CGCTGAAGCC TATCCCGTGC AGTTGCCGGA 50  
pol SHIH TGGCGCTAT AGCCCTCTACA TGGATGACAT CCTGCTGGCC TOC 93

FIG 14

Trad. amores	WKVLPQG	YMDILLAS	
Trad pol SHIH	WKVLPQG-----	AEAYRVQLPDAAYSLYMDILLAS	31
HTLV1 QG-YM	VLPOGFKNSPTLFEMQLAHILOPIRQAFQCTILQYMD		
HTLV2 QG-YM	VLPOGFKNSPTLFEQQLAAVLNPMRKMFPSTTIVQYMD		
HIV1 QG-YM	VLPOGWKGSPAIFQSSMTKILEPFRKQNPEDIVTYQYMD		
HIV2 QG-YM	VLPOGWKGSPAIFQHIMRQVLEPFRKANKDVIQYMD		
MoMulv QG-YM	RLPQGFKNSPTLFDEALHRDLADFRIQHPDLILLIQYVDD		
VISNA QG-YM	VLPOGWKLSPAVYQFTIMQKILRGWIEEHFMIQFGTYMD		
CAEV QG-YM	VLPOGWKLSPSVYQFTIMQEILLEDWIQOHPEIQFGTYMD		
JSRV QG-YM	VLPOGMINSPTLCQKFVATAIAPIVQRFFOLYLVHYMD		
MMTV QG-YM	VLPOGMKNSPTLCQKFVDKAILTVRDKYQDSYTIVHYMD		
MPMV QG-YM	VLPOGMANSPTLCQKYVATAIHKVRHAWKQMYIIHYMD		
HERVK QG-YM	VLPOGMILNSPTICQTFVGRALQPVREKFSDCYIIHYID		
ERV9 QG-YM	VLPOGFRDSPHLFGQALAKDLGHFSS--EGILVQYVDD		
CaMV QG-YM	VPFGLKQAPSIFQRHMDEAFRVFRKFCCV-----YVDD		

## FIG 15

Consensus	TTGGATCCAG TGYTCCCACA GGGCGCTGAA GCGCTATCGCG TGCAGTTGCC	50
Consensus	GGATGCCGCC TATAAGCCTCT ACCTGGATGA CCTSCTGAAG CTTGAG	96

## FIG 16 A

Consensus	TGGATCCAG TGYGCCACA GGGCGCTGAA GCGTATCGCG TGCAGTTGCC	50
B CG13	..... T .....	50
E CG38	..... C .....	50
H CG65	..... C .....	50
Consensus	GGATGCCGCC TATAGCCTCT ACGTGGATGA CCTSCTGAAG CTTGAGTACG	100
B CG13	..... ..G.....	100
E CG38	..... ..C.....	100
H CG65	..... ..C.....	100
Consensus	TGGATGACCT GCTGAAGCTT GAG	123
B CG13	..... ....	123
E CG38	..... ....	123
H CG65	..... ....	123

## FIG 16 B

TGGATCCAG TGYGCCACA GGGCGCTGAA GCGTATCGCG TGCAGTTGCC 50  
 L D P V L P Q G A E A Y R V Q L P

GGATGCCGCC TATAGCCTCT ACGTGGATGA CCTSCTGAAG CTTGAGTACG 100  
 D A A Y S L Y V D D L L K L E Y V

TGGATGACCT GCTGAAGCTT GAG 123  
 D D L L K L E

## FIG 16 C

Consensus	<u>TGGATCCAG TGYGCCACA GGGCGCTGAA GCGTATCGCG TGCAGTTGCC</u>	50
Consensus	GGATGCCGCC TATAGCCTCT ACGTGGATGA CCTSCTGAAG CTTGAGTACG	100
Consensus	<u>TGGATGACCT GCTGAAGCTT GAG</u>	123

## FIG 16 D

VLPQG GAEAYRVQLPDAAYSL YVDD

FIG 17

AA conservés	VLPOG.....E..R.Q.P.....Y.IDD
HTLV1 QG-YM	VLPOQFKNSPTILFEMQLAHILQPIRQAFQCTILQYMD
HTLV2 QG-YM	VLPOQFKNSPTILFEQQLAAVLNPMRKMFP1STTIVQYMD
HIV1 QG-YM	VLPOQWKGSPTAIFQSSMTKILEPFRKONP <del>D</del> IVIYQYMD
HIV2 QG-YM	VLPOQWKGSPTAIFQHTIMRQVLEPFRKANKD <del>V</del> IIIQYMD
MoMulv QG-YM	RLPQGFKNSPTILFDEALHRDLADFRQHQPD <del>L</del> LLQYDD
VISNA QG-YM	VLPOQWKLSPTAVYQFIMQKILRGWIEEH <del>P</del> MIQFGIYMD
CAEV QG-YM	VLPOQWKLSPTSVYQFIMQEIL <del>E</del> DWIQQH <del>P</del> EIQFGIYMD
JSRV QG-YM	VLPOGMINSPTLCQKFVATAIAPVRQRFQ <del>L</del> YLVH <del>Y</del> MD
MMTV QG-YM	VLPOGMKNSP <del>T</del> LCQKFVDKAILTVRD <del>K</del> YQDSYTV <del>Y</del> MD
MPMV QG-YM	VLPOGMANSPTLCQKYVATAIHKVRHAWKQMYII <del>Y</del> MD
HERVK QG-YM	VLPOGMLNSPTICQTFVGRALQPVREKFSDCYII <del>Y</del> MD
ERV9 QG-YM	VLPOQFRDSPHLFGQALAKDLGFSS-- <del>P</del> GT <del>L</del> VLQYDD
TradMSRV-2B 2YV	VLPOG----- <del>A</del> FAYRVOLPDAAYSLYVDD <del>L</del> KLEYVDD
TradMSRV-2B	VLPOG----- <del>A</del> FAYRVOLPDAAYSLYVDD
CaMV QG-YM	VPFGLKQAPSIFQRHMDEAFR <del>V</del> FRKFCCV-----YVDD

FIG 18 A

Consensus	CTTGGATCCA GTGYTROCMC AGGGNGCTGA AGCCTATCGM NTGMAGTINGC	50
MAJ	CTTGGATCCA GTGtTgCCcC AGGGcGCTGA AGCCTATCGc gTGCAGTtGC	50
VAR	ca a a a	50
Consensus	CGGNINCCGN CTWIAGCCNC TACRTGGWIG ACCINNTGAA GNTKGAN	97
MAJ	CGGaTgCCGc CTaTAGCtC TACgTGGaTG ACCTccTGAA GcTtGAg	97
VAR	t a t t g	97
VAR	g	97

FIG 18 B

an	CTTGGATCCA <u>GTGTIGCCCC</u> AGGGCGCTGA AGCCTATCGC GTGCAGITGC	50
prot	<u>V L P O G A E A Y R V Q L P</u>	
an	CGGATGCCGC CTATAGCCTC <u>TACGTGGATG</u> ACCTCCTGAA GCTTCAG	97
prot	D A A Y S L <u>Y V D D</u>	

FIG 19 A

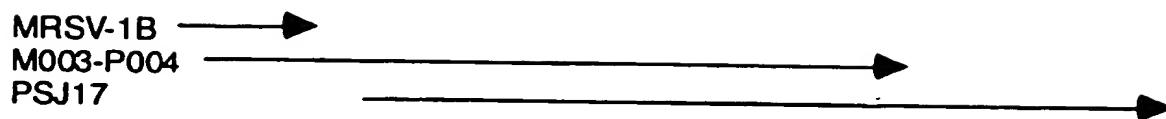
Consensus	TGGAAAGTGY TGCCMCAGGG CGCTGAAGCC TATCGCGTGC AGTTGCGGGA	50
LigT4	-----.....A.....	44
pol SHIH	.....T.....A.....	50
LBpatDut4	-----...T.....C.....	44

Consensus	TCGGCGCTAT AGCCCTCTACR TGGATGACAT CCTGCTGGCC TCC	93
LigT4	.....G.....	72
pol SHIH	.....A.....	93
LBpatDut4	.....G.....	72

FIG 19 B

Consensus	--VLPOQGAEA YRVQLPDAAY SLYVDD---	31
Trad pol SHIH	WK.....	31
Trad LigT4	.....	24
Trad of LBpatDut4	.....	24

FIG 20



PSJ17	-----	-----	-----	
MRSV-1B	GTGTGOCAC	AGGGGTTCAAG	GGATAGCCCC	CATCTATTIG
M003-P004	-----	-----	-----	GGCAGGCCATT
	-----	TCAG	GGATAGCCCC	CATCTATTIG
	-----	-----	GGCAGGCCATT	
				50
				34
PSJ17	-----	-----	-----	
MRSV-1B	AGCCAAGAC	TIGAGOCAGT	TCTCATACCT	GGACACTCTT
M003-P004	AGCCAAGAC	TIGAGOCAT	TCTCATACCT	GGACACTCTT
	-----	-----	-----	GTCCTTCGGT
	-----	-----	-----	100
	-----	-----	-----	84
PSJ17	-----	-----	-----	
MRSV-1B	ACATGGATGA	C-----	-----	CAA
M003-P004	ACATGGATGA	TTCATTTTA	GCTGCCCGTT	CAGAACCTT
	-----	-----	-----	GTCGAAATCAA
	-----	-----	-----	3
	-----	-----	-----	111
	-----	-----	-----	134
PSJ17	GCACACCAAG	AACTCTAAA	TTTCTCACT	ACCTGGCT
MRSV-1B	-----	-----	-----	ACAAGGTTTC
M003-P004	GCACACCAAG	AACTCTAAA	TTTCTCACT	ACCTGGCT
...//...	-----	-----	-----	ACAAGGTTTC
...//...	-----	-----	-----	53
...//...	-----	-----	-----	111
PSJ17	CGAGGGATGA	GCTTGCAACC	CGTGGCATAAC	CTGAATAAGG
MRSV-1B	-----	-----	-----	AAATTGATGT
M003-P004	CGAGGGATGA	GCTTGCAACC	CGTGGCATAAC	CTGAATAAGG
	-----	-----	-----	AAATTGATGT
	-----	-----	-----	503
	-----	-----	-----	111
	-----	-----	-----	634
PSJ17	AGTGGCAAAG	GGTGGCTC	AINGTTATG	GGTAATGGNG
MRSV-1B	-----	-----	-----	GCAGTAGCAG
M003-P004	AGTGGCAAAG	GGTT	-----	-----
	-----	-----	-----	553
	-----	-----	-----	111
	-----	-----	-----	648